Escaping the search box:
Developing digital interfaces that encourage exploration and discovery in cultural heritage collections.

Ben Ennis Butler
2016

A thesis submitted for the degree of
Doctor of Philosophy in Communication
University of Canberra

UNIVERSITY OF
CANBERRA

Abstract

Over the past decade there has been a huge push by cultural institutions to digitise their collection material and facilitate access to it on the Web. In the process of doing this, many galleries and museums have taken their existing catalogue style systems and replicated them online. Most rely on a keyword-based method that is targeted at a user who already knows collections, understands how they are structured and how to search them effectively. There is little room for the non-specialist, the visitor who just wants to have a look around, to browse, explore and play. I ask, is it possible to create a digital collection access interface that allows exploration and discovery to occur?

Taking advantage of new technologies I have created six innovative and experimental interfaces using data from the National Gallery of Australia's Australian Prints + Printmaking website and combined data visualisation techniques with modern web design methods to create new ways of accessing the collection. They emphasise relationships within the collection and provide displays that are denser and richer than conventional Web pages. Results from a mixed-method evaluation study show how the interfaces allow for serendipitous discovery and free-form exploration to occur in ways that are beneficial to the user.

In the development of these rich visual interfaces the research project addresses a number of challenges and issues, from the reliability of data to the limits of the Web browser and traditional evaluation techniques.

The research project demonstrates that it is possible to develop effective Web-based generous interfaces that encourage free-form exploration and serendipitous discovery within digital cultural heritage collections.

Authorship statement

Except where clearly acknowledged in footnotes, quotations and the bibliography, I certify that I am the sole author of the thesis submitted today entitled *Escaping the Search Box. Developing digital interfaces that encourage exploration and discovery in cultural heritage collections.*

I further certify that to the best of my knowledge the thesis contains no material previously published or written by another person except where due reference is made in the text of the thesis.

The material in the thesis has not been the basis of an award of any other degree or diploma except where due reference is made in the text of the thesis.

The thesis complies with University requirements for a thesis as set out in the Examination of Higher Degree by Research Theses Policy.

Refer to:
http://www.canberra.edu.au/current-students/current-research-students/hdr-policy-and-procedures

Ben Ennis Butler 10/07/2016

Sam Hinton 10/07/2016

Acknowledgements

My project would not have been possible without assistance from the Department of Australian Prints and Drawings at the National Gallery of Australia. I give my thanks to them for providing access to the data and their ongoing support of the project.

To the friends of the project, and those I met at conferences along the way, I have appreciated your feedback and thoughtful reflections. I would like to give particular acknowledgement to Marian Dörk, Tim Hart and Jonny Brownbill.

Special acknowledgment must go to my colleagues at the University of Canberra, particularly Tim Sherratt and Erin Hinton, as well as Jen Webb, Tracy Ireland and the Creative and Cultural Practice staff who provided support and encouragement throughout my candidature.

Special thanks also to Ann, Lisa, Sarah and all at The Garden, who kept me sane and provided a wonderful non-computer focused work environment in the early years of my PhD. Also, to Paulie, Prue and the Tilley's staff, I have greatly appreciated your support over the past four and a half years.

Since the start of my research project in August 2011, many of my friends have attempted to understand what I actually do. Despite their uncertainty they have continued to encourage me and provided invaluable friendship and continuous support along the way. To Caito, Georgia, Huey and Darren, Katy and Chris, Kelly, Mimosa and Ken, Matt, Meg, Myles, Miranda, Sophie and Cathy, I give my thanks. Additional, extra special gratitude, goes to Rosa and Beas, Jude and Eliza.

To my supervisory panel, Sam Hinton and Geoff Hinchcliffe, thank you for everything, I appreciate all that you have done for me. To my primary supervisor, Mitchell Whitelaw, thank you for convincing me to go down this path and providing unwavering support and encouragement along the way. I will forever be

grateful of your assistance, generosity, trust and guidance.

I give special thanks to Laura (Soddy), and my extended family, Carolyn, Russell and Harry.

I would not have achieved any of this if it were not for my wonderful family, Helen and Roger, Jack and Amy and Helen Maxwell, who have been there for me at every stage of the process. I am truly thankful for all that they have given me.

Finally, to Rose, for everything.

Online material

Access to the six interfaces and the two websites described in Chapter six is available via the following website:

http://phd.beneb.com

Alternatively, the direct links to the creative works are:

Works and Networks:
http://printsandprintmaking.gov.au/explore/works-and-networks

Decade Summary:
http://printsandprintmaking.gov.au/explore/decade-summary

All Artists:
http://printsandprintmaking.gov.au/explore/all-artists

The Fader:
http://printsandprintmaking.gov.au/explore/the-fader

Subjects Explorer:
http://printsandprintmaking.gov.au/explore/subjects-explorer

Timeline:
http://printsandprintmaking.gov.au/explore/timeline

Table of Contents

List of Figures

List of Graphs

List of Tables

1 Introduction

A small rectangular box.

I could be describing an artwork by an acclaimed Australian artist. Instead, I'm describing the predominant way to access artworks by Australian artists. Next to the small rectangular box, is another small box. This one has a single word in it. It says 'Search'.

When it comes to accessing cultural institutions' rich collections online, the primary method of access is through search-based interfaces. This access technique became dominant in the late 1990s as galleries and museums digitised their collections and replicated their traditional catalogue style systems online. Search can be very effective—provided you know what you are looking for. Such search-based interfaces are targeted at a user who already knows about collections, understands the structure of the data and how to query it effectively. They leave little room for the layperson. What if the user does not know what they want to search for? What if they do not know the collection? In these cases, search does little to encourage a user to explore a rich cultural heritage collection and limits the potential for discovery.

This thesis seeks to demonstrate how digital interfaces can be developed to encourage exploration and discovery in cultural heritage collections. It is a practice-based research project that involved the creation of six innovative Web-based works.

It is a significant project, as the advent of the Internet has brought dramatic changes to the ways we access our cultural heritage collections. We no longer need to visit a physical gallery or to pick up a book to view items from a collection, instead we turn to the Web, where access to digitised collections is available anytime and from any location. The material in these collections, from artworks to books, glass plate negatives to classified documents, has national and international importance as it represents our shared cultural heritage. These collections are integral to the history of our nation and are crucial for the construction of our cultural identity.

The volume of such online collections can be huge: the Victoria and Albert Museum[1] provides access to over 1.1 million items; the Powerhouse Museum[2] offers 116,449 objects, whilst the Tate[3] has more than 70,000 items online. Additionally, there are extensive resources devoted to increasing the wealth of information available, with the National Gallery of Australia, in particular, listing the digitisation of the national art collection as a key priority in their strategic plan (NGA, 2015). The stakes are high, with the Web challenging the ways in which we access, encounter and interact with cultural heritage material in our connected society. As a result, the potential impact of my research project is substantial, as the techniques and methods developed in my work can be applied to all digital cultural heritage collections.

To create my Web-based works I worked exclusively with the Australian Prints collection at the National Gallery of Australia. This vast collection of prints from Australia and the Pacific region has more than 54,000 artworks by over 19,000 artists, making it the most comprehensive collection of its type in the world. Online access is provided through the specialist Australian Prints + Printmaking (AP+P) website hosted by the National Gallery: http://printsandprintmaking.gov.au. In addition to detailed information on artworks and artists in the collection the site provides rich structured data on related exhibitions, galleries and bibliographical references. The wealth of information makes it an invaluable resource for researchers and members of the general public and a perfect collection for my purposes.

During my candidature I created six interfaces: *Works and Networks*; *Decade Summary*; *All Artists*; *The Fader*; *Subjects Explorer* and *Timeline*, each of which provides an alternative method for accessing data from the AP+P collection. The development of these interfaces was informed by my research questions, which revolved around the broad concept of encouraging free-form

1 http://collections.vam.ac.uk
2 http://www.powerhousemuseum.com/collection/database/menu.php
3 http://www.tate.org.uk/about/our-work/collection

exploration and serendipitous discovery in cultural heritage collections. I asked: *"How can digital interfaces that engage users in exploring cultural heritage collections be created? How can data visualisation techniques be developed to encourage exploration and discovery?"* And, further, *"How can new Web-based technologies be leveraged to create exploratory digital interfaces?"*

The research project sits within a number of fields of contemporary cultural practice. It combines data visualisation techniques with modern Web design methods in order to create novel ways of accessing digital cultural heritage collections. From a cultural heritage perspective, my project is aligned with the 'musetech' domain—a specific professional audience interested the application of technology in cultural institutions.

This Exegesis presents an account of the inquiry undertaken in the practice-based research project. It discusses the theoretical and practical contexts in which the creative work is placed, and identifies a need for the creation of more engaging collection access techniques that are not reliant on the conventional search box. It demonstrates the new knowledge that emerges through the production, evaluation and critical reflection of the creative works.

1.1 Chapter outline

Chapter two outlines the practical context in which the creative work is situated. It introduces a survey of online collection access techniques from a number of cultural institution websites, which shows the dominance of keyword-based search as the primary point of access. I outline the emergence of the Web and discuss how galleries and museums experimented with various techniques, including virtual tours and CD-ROM based interfaces, which provided early digital access to their collections. I introduce the arrival of the first search-based interface in 1993 and identify other leading early works, including The Fine Arts Museums of San Francisco *The Thinker* website. Following this, I outline a number of collection access case studies which start to investigate the possibilities of

creating more engaging methods of collection access. These include early experimental works, such as SFMOMA's *ArtScope* (2009), and more recent websites by the Walker Art Centre (2011), Rijksmuseum (2012) and Cooper Hewitt (2016), which disrupt traditional methods of online access by departing from the search box and emphasising visual and conceptual exploration.

Chapter three considers a number of key theories and concepts that have provided a theoretical framework for the project. I begin by discussing the concept of information retrieval (Mooers, 1951) before describing how Taylor's (1962) theory of information needs is integral to a user's ability to form an effective search query. I demonstrate that relying on search is problematic, as the user is often unable to specify an appropriate query or might not know what they are looking for. In order to overcome these issues, I introduce the notion of information seeking, which provides conceptual models to describe a user's process of seeking information. I describe how task-based models, such as those by Kirkelas (1983) and Kuhlthau (1991), do not reflect our everyday information seeking experiences and instead consider how alternative models, such as exploratory search (Marchionini, 2006), browse (Bates, 2007; Case, 2012), berrypicking (Bates, 1989) and the information flaneur (Dörk et al., 2011) are more appropriate in an exploratory context. I discuss visual information seeking and describe how Shneiderman's (1996) Visual Information Seeking mantra provides a foundation for the creative work in my project. Additionally, I outline Dörk et al.'s (2012) "visual information exploration" as a useful Web-centric model to consider. The next section identifies theories of digital cultural heritage collection interfaces, in particular, Whitelaw's (2012) notion of generous interfaces, the idea of overview (Greene et al., 2000) and focus+context displays (Card et al., 1999). Finally, I describe the background of data visualisation and introduce the concept of cultural analytics (Manovich, 2007), before reflecting on the application of data visualisation techniques within the digital humanities (Drucker, 2013).

Chapter four introduces the research methodology that informed my project. To answer my research questions effectively I combined my practice-based methodological approach with a number of methods. These were selected as they allowed me to bring together both theoretical and practical components. I will refer to Sullivan's braid metaphor (2005), sketching in code, primary and secondary research, prototyping tools and mixed-method evaluation Finally, I present a readaptation of the braid to illustrate how it encapsulates the research process I have employed in the project.

Chapter five elaborates on the production of the six creative works. I discuss the development of each interface, its aims and design intentions before describing the outcome, specific features and technical considerations. As the project progressed I became more adept at implementing complex solutions and experimenting with various Web design and data visualisation techniques. I show how each interface reveals different aspects of the collection and how, as a group, they create multiple points of entry that give the user wide-ranging opportunities for exploration and discovery. The first three interfaces, *Works and Networks*, *Decade Summary* and *All Artists,* were produced at the start of the research project between 2011 and early 2013 in collaboration with my supervisor Dr Mitchell Whitelaw. *The Fader* followed in 2013 and my most complex works, *Subjects Explorer* and *Timeline*, were undertaken during 2015.

Chapter six presents the evaluation component of the research project. This was a two stage process, the first of which was completed after the production of *Works and Networks*, *Decade Summary* and *All Artists*. It involved evaluating the effectiveness of these interfaces in encouraging free-form exploration and serendipitous discovery. I used a mixed-method evaluation study that consisted of think-aloud observations made by selected participants, an open Web survey and data logging. The results of the study revealed that users regarded the interfaces as effective collection access tools, and significantly, that they could successfully apply them to engage in the process of free-form exploration and serendipitous discovery. The completion of this study was a formative component of my research project, as

the results directly informed the development of *Subjects Explorer* and *Timeline*. The second stage of the evaluation process considers critical feedback and evaluation of the final interfaces I produced.

Chapter seven outlines my contributions to new knowledge and elaborates further on the research questions that governed my project. My contributions to knowledge are as such: I have created six unique interfaces which allow visual exploration of a large digital cultural heritage collection; I have developed new techniques for promoting free-form exploration and serendipitous exploration by combining methods from data visualisation and modern Web design, which are applicable to any other collection of cultural data; and I have evaluated the works and demonstrated how a mixed-method research study can provide valuable insights into this style of work. I address the research questions by reflecting on the themes of development and exploration.

Finally, in Chapter eight, I summarise the project, the methods and techniques I have developed, and the outcomes achieved over the course of my PhD. I conclude by identifying the contribution my research has made to the field of cultural heritage through the creation of knowledge about how to develop interfaces which encourage exploration and discovery in Web-based collections.

This research project has shown there are substantial limitations in how access is provided in digital cultural heritage collections. While eminent public collections hold objects of national and international importance, most galleries and museums continue to provide online access to them only via keyword-based search interfaces. My research project provides unprecedented new access opportunities through the demonstration of viable alternatives to the traditional search-based approach. I have used real collection data to create new tools, methods and techniques that are applicable to any other collection of cultural data.

2 Practice Review

2.1 Introduction

In this chapter I will discuss the practical field in which my research project is situated. In the following chapter I will review relevant literature in order to provide a theoretical framework for the project.

I will begin by presenting a survey of online collection access across 23 cultural institution websites which shows the dominance of search-based interfaces as the primary point of access. Having established the current state of play, I look back to the emergence of the Web to understand how galleries and museums responded to the new medium. I discuss how novel techniques, such as virtual tours and the CD-ROM-based Micro Gallery provided users with early digital access to cultural heritage collections. I then introduce the first gallery and museum websites which gave online access to their collections, before reflecting on a number of collection access tools created after 2000. These tools move away from a reliance on search and start to investigate the possibilities of developing more engaging methods of collection access. In particular, I discuss: *ArtScope*; *Preservation of Favoured Traces*; *Visible Archive*; the Walker Art Center collection interface; Google Art Project; Circus Oz: *The Living Archive*; the Rijksmuseum and finally, the Cooper Hewitt collection interface. These examples provide a contextual background for my work, and show how it sits in a rich field of practice. In line with the rationale of this project my review concentrates on Web-based interfaces rather than stand-alone programs.

In the discussion that follows, I refer to a number of theories which I examine in detail in the second part of this chapter. These are: Ben Shneiderman's Visual Information Seeking mantra, Whitelaw's concept of generous interfaces and Marian Dörk and collaborator's *information flaneur*. As a way of providing context for this analysis, I will briefly introduce some of these now.

Shneiderman's (1996) Visual Information Seeking mantra aims to describe the process users engage in when seeking information in interactive environments. Shneiderman outlines his "mantra" for visual exploration as "overview first, zoom and filter, then details-on-demand" (Shneiderman, 2006). He sees the overview as the first step, giving the user a 'bird's-eye' view of the information; a user can then zoom and filter as they focus their interest; before receiving detailed information on demand. Additional steps are "relate", "history" and "extract". Whitelaw's concept of generous interfaces introduces a number of guidelines for the design of digital collection interfaces. The principles are centered around the notion of 'generosity', and are explained in detail later in this chapter. Dörk et al.'s (2011) information flaneur is a visual information seeking model which emphasises exploration over traditional task-based models.

Table 1
Survey of collection access interfaces (2010, 2011, 2016).

Key	
STC	Straight To Collection
C	Combined
H	Highlights
S	Search
X	Available on website

2.2 Survey of collection interfaces

In 2010, 2011 and 2016 I conducted a survey of 25 cultural institution websites and analysed how they provided online access to their collections. In 2010 and 2011 I found that every site in Australia and a vast majority of those overseas, relied on a search-based interface as the primary access point to the collection.

Revisiting these sites in 2016 revealed that search-based interfaces were still the most popular entry point to a collection. In some cases, the terminology used to access this information had changed, for example, from 'Search the collection' to 'Explore'. However, the search box invariably remains.

Most websites from 2016 included a section containing 'highlights' of the collection; see for example, the NGA and AGNSW websites. These are curated selections of artworks presented in slightly more appealing ways than a standard search results page and are targeted at users who are unfamiliar with the collection. However, the static representations showcase a particular, predetermined slice of the collection rather than attempting to encourage freeform exploration and discovery.

	2010				2011				2016			
	Primary access	Highlights	Search	Facets	Primary access	Highlights	Search	Facets	Primary access	Highlights	Search	Facets
NLA	S	N	X	X	Unchanged from 2010				H			N
NGA	S	N	X		S		X		S	X	X	
AGNSW	S	X	X	X	Unchanged from 2010				H	X	X	X
SLV	S		X		S		X		S		X	X
NGV	S	X	X		C	X	X		C	X	X	X
QAG	H	X			Unchanged from 2010				H	X	X	X
AGWA	S	X	X		Unchanged from 2010				Unchanged from 2010			
AGSA	H	X	X		Unchanged from 2010				Unchanged from 2010			
NPG	S		X		Unchanged from 2010				STC		X	X
Powerhouse Museum	S	X	X	X	Unchanged from 2010				Unchanged from 2010			
SFMOMA	S		X		Unchanged from 2010				STC	X	X	X
MOMA	C	X	X	X	Unchanged from 2010				STC		X	X
Guggenheim	H				Unchanged from 2010				C	X	X	X
Getty Gallery	C	X	X	X	C	X	X	X	C	X	X	X
NGA (US)	C	X	X	X	Unchanged from 2010				C	X	X	X
Victoria & Albert	S		X	X	Unchanged from 2010				C	X	X	X
Te Papa	C	X	X	X	Unchanged from 2010				C	X	X	X
British Museum	C	X	X	X	C	X	X	X	C	X	X	X
Louvre	S	X	X		Unchanged from 2010				Unchanged from 2010			
Digital NZ	S		X	X	Unchanged from 2010				C	X	X	X
National Gallery (UK)	H	X		X	Unchanged from 2010				Unchanged from 2010			
Tate	S	X	X	X	Unchanged from 2010				STC		X	X
Cooper Hewitt	H	X	X	X	Unchanged from 2010				STC	X	X	X

NPG: National Portrait Gallery
NLA: National Library of Australia
AGSA: Art Gallery of South Australia
AGWA: Art Gallery of Western Australia
NGA (US): National Gallery of America
MOMA: Museum of Modern Art New York

The reliance on search-based interfaces for discussing large digital cultural heritage collections seems to contrast to their open and inclusive nature. Visiting a cultural institution can be an exciting experience: you walk through the front doors, and into gallery spaces full of artworks and objects you may have never seen before. The architecture of the building seeks to engage you; the gallery spaces and strategic lighting are cleverly designed to encourage you to wander through them; the gift shop draws you in to browse and make purchases. The whole process culminates in an experience to remember and reflect on.

As evidenced by websites such as those from the NGA, Getty Museum and the NGA (US), there appears to be a continuing failure on the part of some cultural institutions to embrace the possibilities of the Web as a platform for accessing and representing their collections. A majority of websites either attempt to emulate a physical visit or provide an uninspired reproduction of the back-of-house catalogue system. Often little thought is given to leveraging the power of the Web to enhance access to and experiences of works of art.

2.3 Early Development

The World Wide Web is an ubiquitous part of contemporary life even though its development is relatively recent, dating from the late 1980s.

The World Wide Web[1] originated as a technology that enabled researchers to easily share information with each other (Cailliau, 2000). It was created in March 1989 by Tim Berners-Lee, a software engineer at CERN in Switzerland (W3C, 2015b). In April 1993 two ground-breaking events laid the foundations for the future of the Web as we know it. On April 22, the Mosaic Web browser was released. It was the first to include a graphical user interface

1 The internet and the World Wide Web are not the same. The internet is the protocol that allows the particular style of communication to occur between connected computers, the Web is one platform that uses the internet, examples of other platforms include email, VoIP and FTP. Hereafter I refer to the World Wide Web as just the Web.

and transformed how a user could access and view Webpages. It had an image tag which allowed text and images to be displayed alongside each other. Previously, images and other multimedia were hidden behind a textual link which would open the resource in a new window (Calore, 2010). Writing in *Wired*, Gary Wolfe described the excitement that surrounded the browser's release:

> With Mosaic, the online world appears to be a vast, interconnected universe of information. You can enter at any point and begin to wander; no internet addresses or keyboard commands are necessary. (Wolfe, 1994)

Quickly following the release of Mosaic came the announcement by Berners-Lee on 30 April that the technology behind the Web would be released to the public on a royalty free basis (Cailliau, 2000). Berners-Lee explains that if the "technology had been proprietary... it would probably not have taken off. The decision to make the Web an open system was necessary for it to be universal" (Berners-Lee, 1998). The World Wide Web Consortium (W3C) claims the commitment to being non-proprietary was a turning point in the development of the Web because it "sparked a global wave of creativity, collaboration and innovation never seen before" (W3Cb, 2015). Berners-Lee recognised the importance of sharing the technology in order to encourage growth and development, and consequently at this early stage of the Web's evolution, ideas about generosity began to emerge. The concept of generosity underpins Whitelaw's notion of generous interfaces, which I will elaborate upon in Chapter three.

One of the earliest voices in the then emerging field of online museums was Jonathan Bowen, a research officer at the Oxford University Computing Laboratory with an interest in museums and access to the Web. He created the first directory of museum websites, which was added to Berners-Lee's Virtual Library[2] in June 1994 (Bowen, 2010). At that point in time, Bowen explained how:

> the main museums with early websites tended to be those with academic associations since this gave them relatively easy and cheap access to the Internet and to the expertise needed to install and run a webserver. (Bowen, 2010)

Bowen gave three examples where this occurred: in the USA, the University of California's Museum of Palaeontology; and in the United Kingdom, the Natural History Museum and the National Museum of Science and Industry.

In 1995, at the third International Conference on Hypermedia and Interactivity in Museums, David Bearman (1995), described how the focus of the conference had shifted from discussions about CD-ROM publication to an emphasis on "remote access to interactive multimedia, made possible… by the explosion of the World Wide Web which didn't even exist when we held our meeting in 1993." Dixon and Barata (1999) noted that by the late 1990s there were "quite a few art museums on the Web" but many of the sites contained only basic information such as location of the museums and a list of current exhibitions; Suzanne Keene (1998) termed these sites "brochureware".

The increasing awareness of the potential of the Web and its rapid growth were evident as a number of prominent museum websites came online between 1995 and 1997 (see Figure 1).

Maria Piacente (1996) identified three clear themes that emerged in the development of these early museum websites. First was the "Electronic Brochure", the basic brochure style website also identified by Dixon, Barata and Keene (1998); second, the

2 Domain names were mostly unknown at this point, so the Virtual Library was created as a way to provide a directory of content on the web. The library consisted of links to various sites, grouped by subject headings.

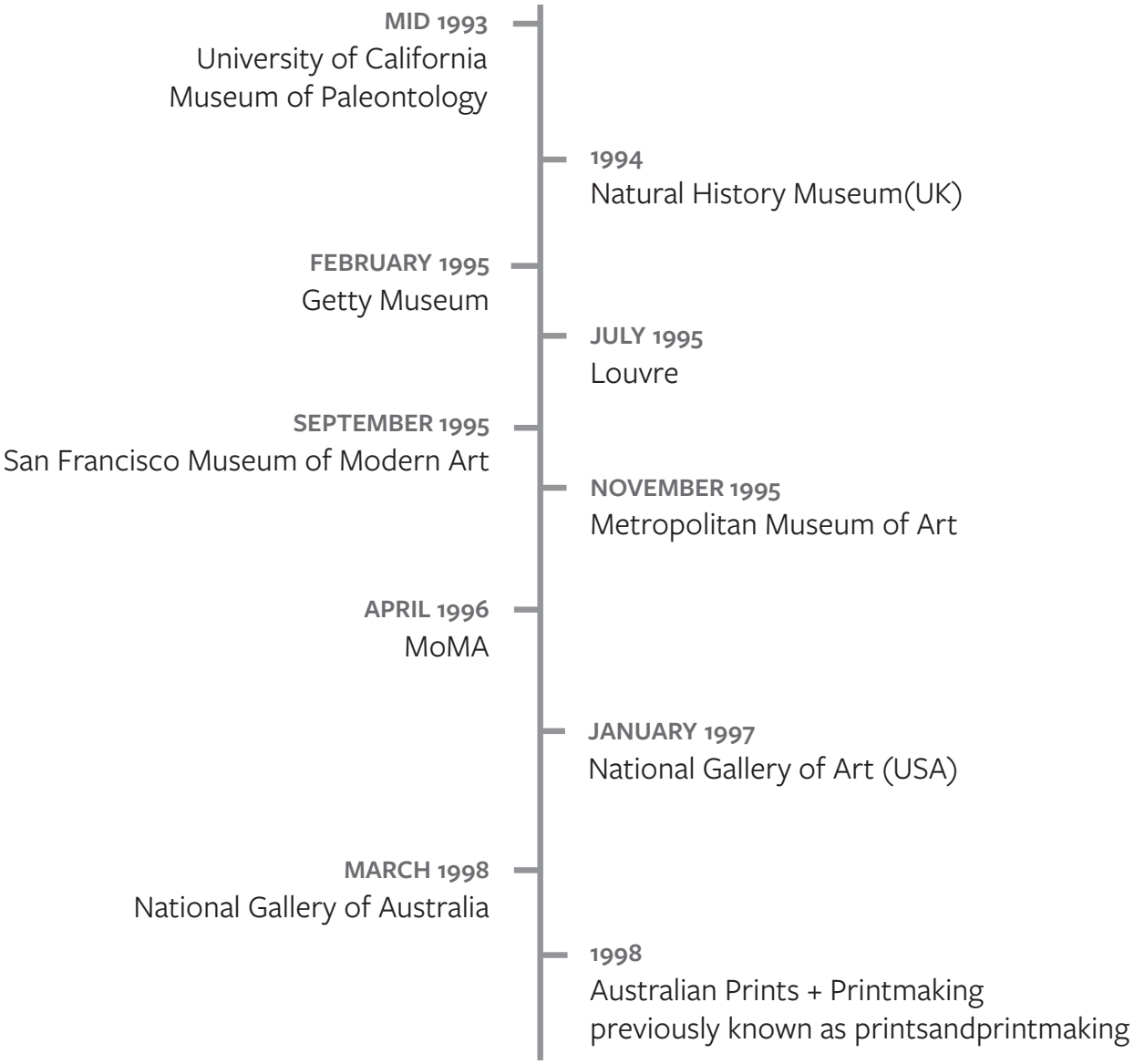


Figure 1
Timeline of early gallery and museum websites

"Museum in the Virtual World" emulated the physical experience of a visiting a museum by including "virtual tours" (Teather and Wilhelm, 1999) or "ready-made displays sometimes called Web museums" (Nilsson, 1997); and finally, the "True Interactives" were those sites that maintain a connection to the museum and embrace the possibilities of the new digital technology. Sarah Kenderdine (1996) expands upon the latter using the term the "virtual museum."[3] She explains how a traditional museum has a number of functions (education, research, marketing, publication and so on), and a virtual museum will always be embedded within these functions, "however, it can also extend the breadth of information on offer

3 Referring explicitly to the website of an actual museum

to present all collections, and allow multiple interpretations and perspectives." In other words, a museum can reproduce their traditional functions online or they can complement them and build new ones as "facilitated by the new medium" (Teather and Wilhelm, 1999). This could include online access to artworks not on display, particularly important given that over 95% of artworks in most galleries are in storage[4] (Molineux, 2014).

In 1997 the groundbreaking first *Museums and the Web* conference, organised by Archives & Museums Informatics[5], was held in Los Angeles and by the end of the year it was clear that the Web revolution was well on its way. An estimated 650,000 websites were online globally and that number was increasing rapidly (Gray, 1997).

As more websites and users came online, museums started to realise the potential of the Web as a new medium (Bearman & Trant, 1999) with many new possibilities. Bowen writes:

> The Web is a different medium with its own strengths and weaknesses, which should be exploited to enhance the virtual visitor experience. Vast amounts of information such as museum catalogue databases can be made available online at relatively little cost. (Bowen, 1999)

This point about cost is emphasised by a number of researchers, including Paul Kahn, Krzysztof Lenk and Magdalena Kasman (1997). They described how the ability to present a large quantity of detailed collection data on the Web appealed to galleries and museums due to its inexpensive and flexible nature. In their paper *New Web-based Interfaces to Old Databases*, presented at *Museums and the Web* in 1998, Stephen Toney and Karen Donoghue (1998) explain how the Web provides an opportunity to make pre-existing cataloguing databases available at low cost to a worldwide audience, quickly removing traditional geographical boundaries and offering a convenient global point of access around the clock. Theirs was a view supported by Proença et al. (1998) who write that "new multimedia

4 At the NGA in 2014-2015, 1.38% (2405) of the collection was on display
5 http://www.archimuse.com/conferences/mw.html

and interactive technologies represent new ways museums can improve the communication with the public, to attract remote (virtual) visitors... and explore new potentials – some only possible due to recent information technology." It was clear that the rapidly changing technological shift that occurred in the 1990s would have enormous implications for museums and galleries as they navigated the shift from being purely physical to supporting substantial digital engagement.

2.4 Museums in the Virtual World

As previously outlined, many museum websites attempted to reproduce the physical museum experience on the Web. The most popular methods were either through a virtual tour or a virtual gallery experience.

Kahn, Lenk and Kasman (1997) describe how a virtual tour allows visitors who are "geographically remote from the publishing institution" to view an exhibition on the Web. An early example of this was the virtual version of the special exhibition *The Measurers: a Flemish Image of Mathematics in the Sixteenth Century* developed by the University of Oxford's Museum of the History of Science in 1995. Jim Bennett explains how a 'virtual' version can contain much more information than could be presented in a traditional exhibition format, and since the "images are offered at different resolutions, it is often possible to see more detail on the computer screen than in the showcase" (Bennett in Bowen et al., 1998).

In comparison, a virtual gallery attempts to portray the actual physical experience of visiting the gallery, through the use of interactive floor and room maps and technology such as Apple's QuickTime interactive panoramas[6] (Kahn, Lenk and Kasman, 1997). Nilsson (1997) argues that this style of pre-produced online content was, in many cases, "outright boring" and far removed from the experience of a "good physical museum where almost anyone can stroll around and find something of interest."

6 Virtual Galleries still exist—Google Art Project offers such a service

2.5 True Interactives

An early interactive collection access technique was the Micro Gallery developed for the National Gallery in London in July 1991. The first iteration was only accessible by visiting the gallery where it ran on a number of touch screen kiosks installed within the gallery itself. Alex Morrison (1995) describes the concept as follows: "[the aim is to] provide visitors to a museum or gallery with an enjoyable information system that they can use to find out more about the works on display." Images and information on all 2200 artworks of art in the collection were available through the Micro Gallery. The CD-ROM version was released in 1993 and contained the entire contents of the London Micro Gallery with additional guided tours and search facilities (Matthews, 1999; Dixon and Barata, 1999). By 1999 it had sold more than 300,000 copies (Matthews, 1999) indicating that there was significant demand for a product with comprehensive information.

Dixon and Barata (1999) explain that the Micro Gallery provided a popular model that many museums wanted to emulate as it included all the artworks on exhibit, and gave access to images, textual descriptions, artists' biographies and maps of the galleries all accessible via an interface that, by 1993, allowed search (Matthews, 1999). It was therefore an example in which the user experience was being enriched by the use of new technology.

2.6 Early collection access online

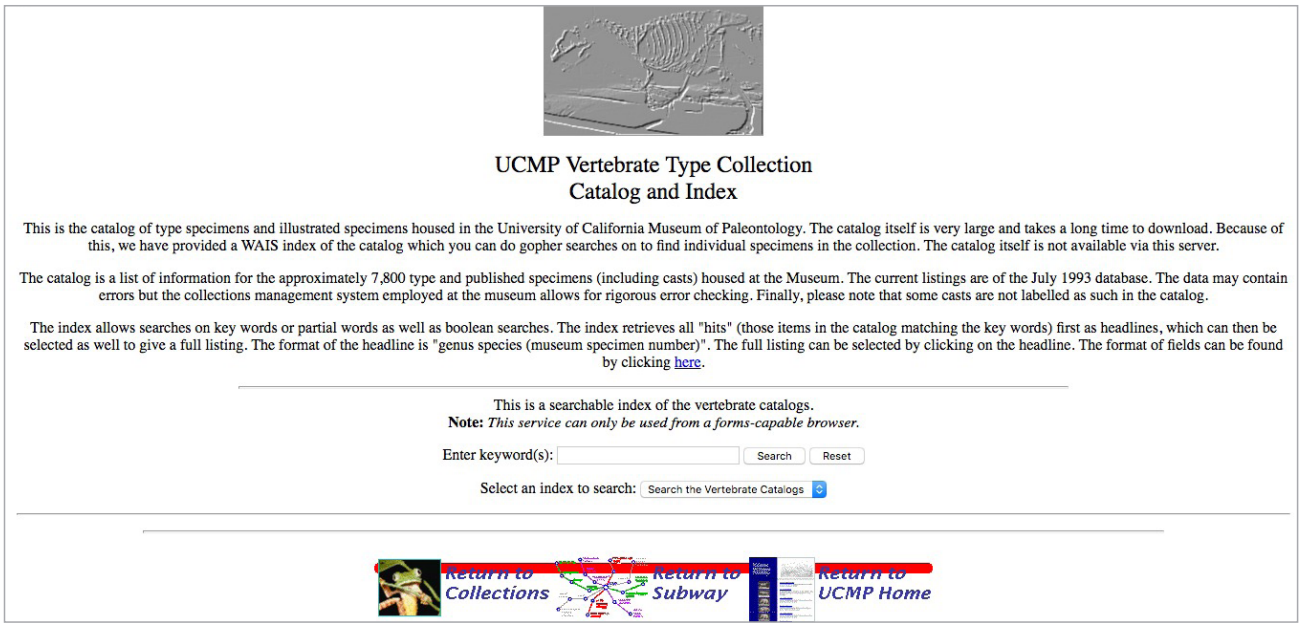
As explained earlier in this chapter, the University of California's Museum of Paleontology launched their ground-breaking first website at some point in 1993, which appears to be the first instance of search in a Web-based digital collection (Figure 2). It included the functionality to search the museum's collection database online (Guralnick, 1997; Bowen, 2010). Robert Guralnick (1997), its creator, explained that "although not elegant, [the collections interface] was pragmatic in 1993 when the kinds of interfaces necessary had yet to be built." Guralnick (1997) wrote that: "in museums that have collections, like many science museums, the Web will become the

standard interface to collections information." A prescient claim, as it would be at least two more years before other museum and gallery websites started to place more comprehensive collection material online (Bowen, 2010; Burton Jones, 2012).

For the period 1994 to 1998, Katherine Burton Jones (2012) identified a number of pioneering museums that established a strong online presence by putting digital collection databases online. Joining the UCMP was the Fine Arts Museums of San Francisco with their *The Thinker* database (Figure 3) which launched in November 1996 and gave access to a collection of more than 70,000 imaged art objects (Talagala et al, 1998); the United States Holocaust Memorial Museum (Figure 4) and the Smithsonian National Museum of American Art HELIOS American photography website that both launched a year later. Examples outside the US include a website developed by the French Ministry of Culture that offered online access to a database of more than 130,000 paintings (Mannoni, 1997).

None of these early sites are still online, however, archived versions are accessible via the Internet Archive. My analysis of these sites show their reliance on search as the primary method of access. For instance, on the UCMP collection catalogue page, the search box is the dominant feature on the page, immediately below which is a long description about the contents of the

Figure 2
Screenshot of UCMP search page from the Internet Archive



collection and instructions about how to query it effectively. Unfortunately search result pages are not captured by the Internet Archive, however given the appearance of the first page, I suspect they would have been returned as a standard list of results.

The Thinker also provides access through search, although in their interface it is also possible to search by the artist's last name, country or by selecting from a drop-down box containing a list of time periods (for example, 18th century). It is unclear whether the time period selection requires a query to be entered, or whether it can return results independently. As with the UCMP site, it is not possible to see the search results, however Asami and Patterson (2000) explain how images within the site were presented using GridPix, a "web-based image presentation system that allows users to zoom in

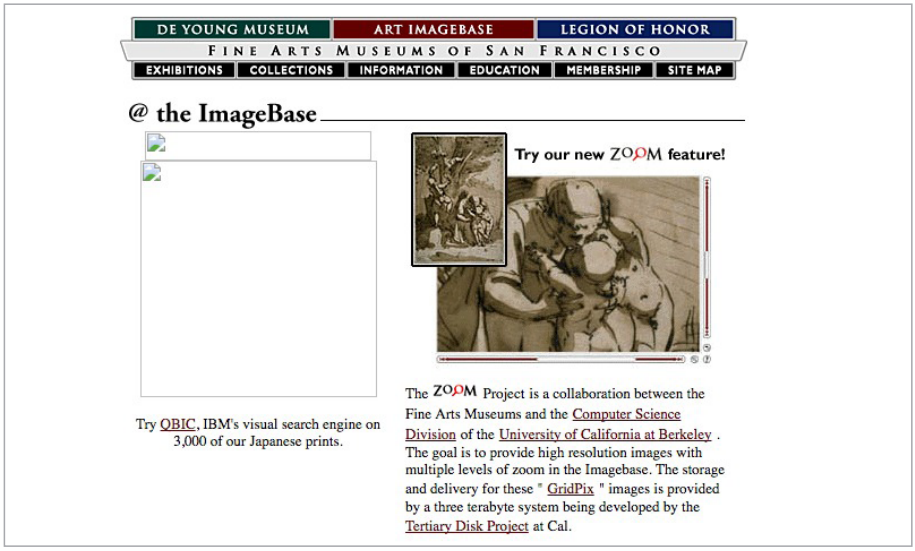


Figure 3
Screenshot of *The Thinker* from the Internet Archive

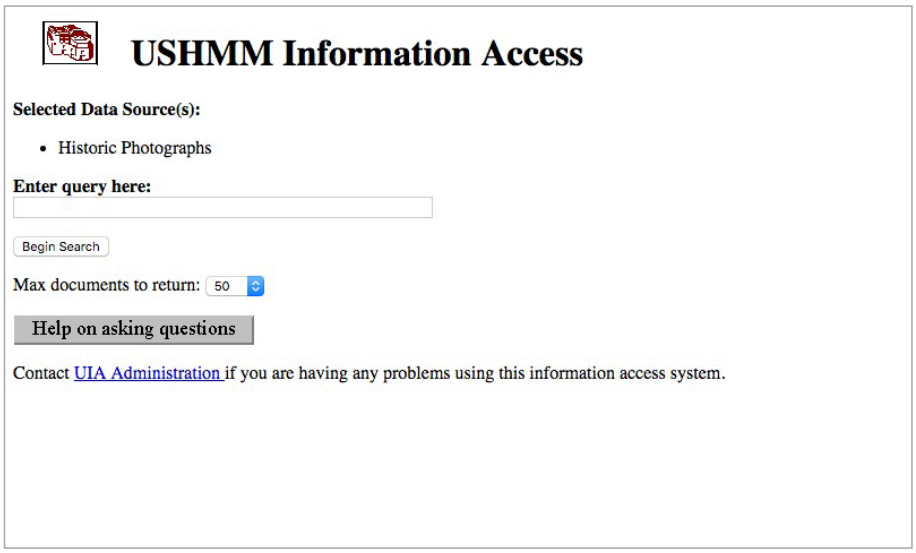


Figure 4
Screenshot of search interface from United States Holocaust Memorial Museum from the Internet Archive

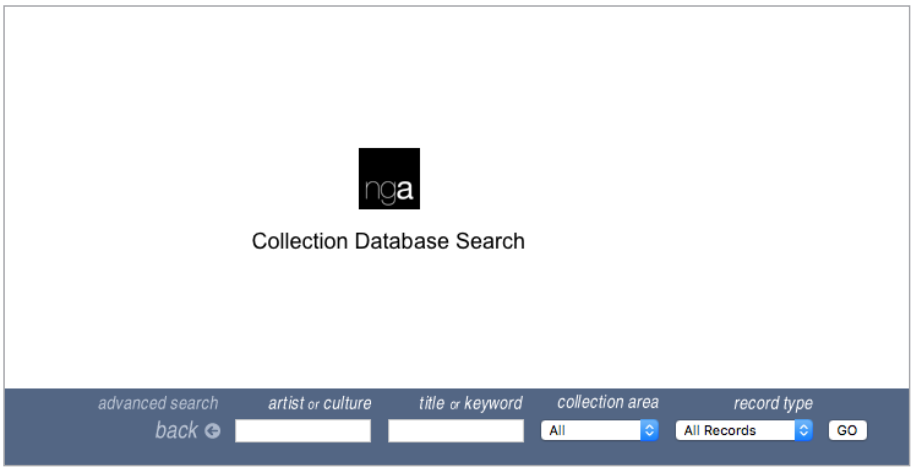


Figure 5
Screenshot of the NGA search page in 2000 from the Internet Archive

to portions of a large image and scroll around to view different parts by presenting the image as a layered, tiled collection of grids." The size of images available online was quite large, up to 6000 by 4000 pixels[7], and this technique allowed access to them without having to download unnecessarily large files. Many of the methods employed in GridPix were later replicated in digital mapping services.

In Australia the creation of a website for the NGA was overseen by Senior Curator of Australian Prints and Drawings, Roger Butler, and Head of Education, Peter Naumann, and was launched on March 6 1998. It included a basic collection search functionality (Figure 5).

Butler recognised the potential of the Web as a way of providing new forms of access and created the Printsandprintmaking website in 1998 as an offshoot of the NGA's main website. The aim was to "provide succinct, reliable information relating to prints, posters and book arts in Australia and the Australasian region" (Butler, 2011), and display it online utilising new access techniques and descriptive cataloguing. It was supported by the Gordon Darling Australasian Print Fund. Printsandprintmaking was the winner of the best research site at the *Museums and the Web* 'Best of the Web' awards in 2008 (Figure 6). It was rebuilt in 2011 and renamed Australian Prints + Printmaking.

The Thinker and UCMP sites both won the inaugural *Museums and the Web* 'Best Research Site' award in 1997 (MW, 1997), and *The*

7 Sample of the quality of image returned here: http://web.archive.org/web/19990127160823/http://now.cs.berkeley.edu/Td/GridPix/

Thinker won again in 1998 (MW, 1998). Butler (2013) and Bowen (1999) argue that these early websites with online catalogues were developed primarily for scholars and researchers who were interested in museums and online access. MacDonald (2015) outlines how such users wanted convenient access to collection data without having to be physically present. The naming of the *Museums and the Web* award category as the 'Best Research Site' reinforces this.

Throughout this period, the nature of the Web was significantly changing. There was a massive increase in connected users, from around 9 million users in 1993, to 119 million users in 1997 and by the turn of the century, to more than 400 million (Roser, 2014). At the same time governments and businesses started to move resources online and many pioneering Web-based companies were launched, including Yahoo in 1994; Amazon and eBay in 1995 and Google in 1998 (Mayo and Newcomb, 2009). The tremendous increase in users and resources available online had the effect of legitimising the platform, it was no longer just a medium used by academics and scientists. The rise of Google and other search engines cemented the popularity of the search box as a method for accessing information online.

Whilst the search engines were providing users with access to information in ways never seen before, there was an implicit assumption within cultural institutions that those accessing the online collections were professional users who already had specialist

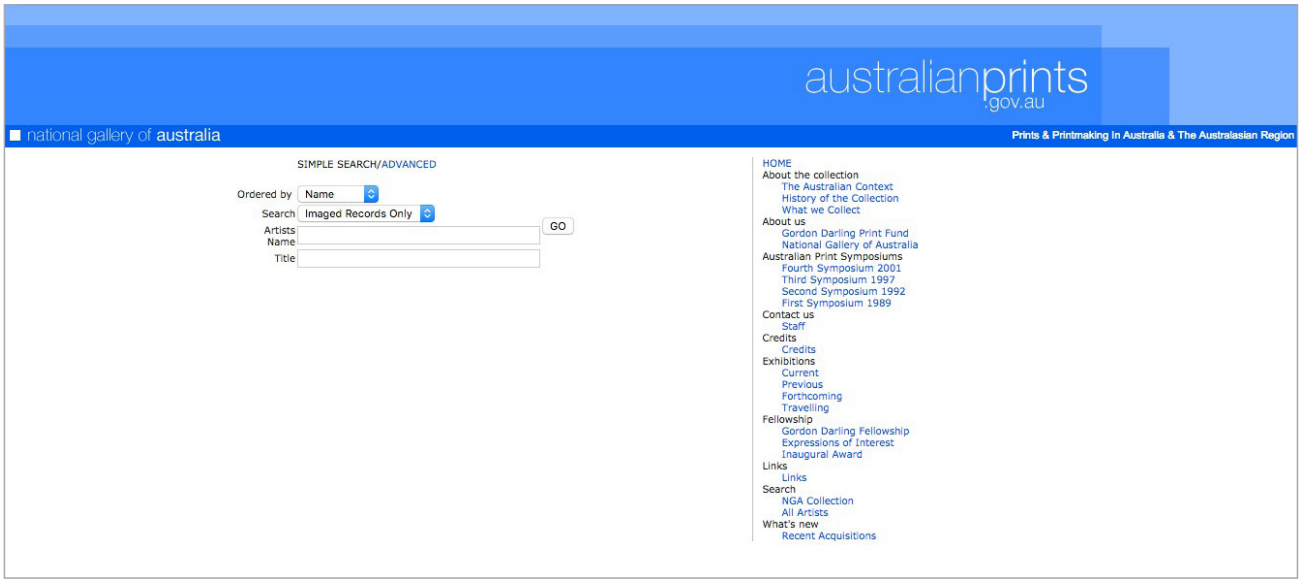


Figure 6
Screenshot of the first printsandprintmaking site from the Internet Archive

knowledge of the collection, or understood the style of language used to catalogue the artworks, and were therefore able to query it effectively. However, Butler (2013) argues that in a connected society, users accessing these online databases could come from anywhere and might not necessarily have a background in, or familiarity with, the arts. In other words, potential users might not just come from a different country but could be referred from a search results page rather than following the process dictated by the gallery website.

2.7 Collection access case studies

I will now consider a number of collection access tools created in the new millennium that have evolved from early access techniques. As with the previous discussion, what follows is not intended as a comprehensive history of the field. Some of the earliest works in this area include: George LeGrady's (2005) visualisations of activity within the Seattle Central Library and Jeanne Kramer-Smyth's (2007) *ArchivesZ* that focused on the scope and availability of records within archival collections. However my aim is to draw attention to some of the exemplary works produced globally which show the potential for new methods of online collection access. I will therefore analyse *ArtScope*; *Preservation of Favoured Traces*; *Visible Archive*; *commonsExplorer*; the Walker Art Center collection interface; Google Art Project; Circus Oz: *The Living Archive*; the Rijksmuseum and finally, the Cooper Hewitt collections interface.

I refer to these works as 'tools', rather than interfaces, as some are stand-alone programs, with live components, rather than all being strictly Web-based interfaces. Significantly they start to investigate the possibilities afforded by the new medium and borrow techniques from other fields, including data visualisation and digital mapping. Some of these tools are developed by the cultural institutions themselves, for example, Rijksmuseum and Cooper Hewitt, whereas others, including *commonsExplorer*, *Visible Archive* and *Preservation of Favoured Traces*, have been created by developers external to the institution. In the latter cases, they come to the project having being influenced by different

cultural perspectives, backgrounds, interests and aims.

ArtScope

ArtScope (SFMOMA, 2009) is one of the most outstanding early examples of an effective visual exploration interface for digital cultural heritage data. Produced in 2009 by Stamen Design for the San Francisco Museum of Modern Art, it is an immersive browser-based experience which allows a user to view and explore 6491 artworks from the SFMOMA collection on one screen (Figure 7). Stamen Design's motto is "show everything" (Jones, 2009) and is clearly obvious in this eye-catching, and radically different, method of online collection access.

In *ArtScope*, the artworks are initially presented through a bird's-eye overview, as a huge grid of thumbnails; a user can then engage in Shneiderman's Visual Information Seeking mantra as they click on the artworks that appeal to them and literally zoom in to focus on them, while a pane on the right of the screen provides further details on demand. Stamen Design is an agency which focuses on innovative digital mapping and they explain how they built *ArtScope* using their own "Modest Maps library, an open-source toolkit that allows designers and developers to quickly build and design zoom-and-pan maps" (Stamen Design, n.d). The zoom-and-pan technique artworks well, as the user can move around the interface at every stage, from the overview to the detailed object view. It is a playful technique which encourages users to engage with the diverse collection as they spot objects of interest. Upon zooming in, the interface loads a higher resolution version of the current view, thus ensuring the display updates quickly and reducing the load time of the whole interface. The cleverly designed interface employs common visual cues to avoid the need for further instructions, for instance, the metaphor of a lens to magnify individual artworks. Visual devices such as this and the familiarity of the online mapping technique facilitate immediate understandings of how to engage with the interface and the collection.

The detailed Information within the right pane provides further

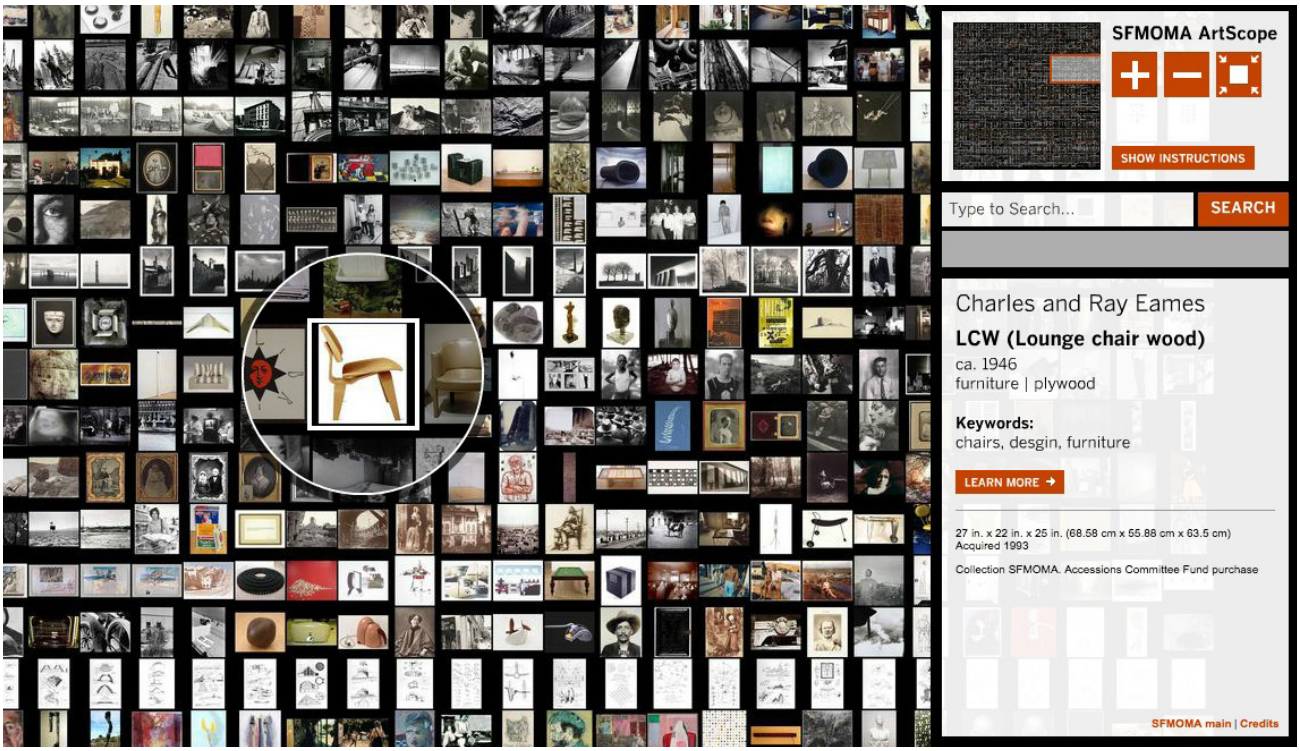


Figure 7
SFMOMA's *ArtScope*

context about the object in focus and the ability to navigate through the display by clicking on the artist's name, artwork date, type or keyword. Applying this method of navigation does not filter the display but instead allows the user to jump from artwork to artwork with the arrow keys or by clicking the 'next' button. This removes the ability to gain a further overview of a subset of the data. For example, by clicking the keyword 'chairs', a label updates to say there are 159 different images, however the display still shows all 6491 artworks. When zooming out to the original overview, there is no indication where those 159 chair images are in the grid of artworks. If the interface were able to be filtered to only show these artworks it would provide the user with an immediate understanding of their focused selection. While *ArtScope* steps away from search, it does not remove it completely. The search field remains in the right corner of the screen and functions in the same way as the keyword example described above, this allows those users who are looking for something particular to still use the interface effectively.

ArtScope was built with Adobe Flash, which whilst extremely popular at the time, is no longer supported on many modern devices and Web browsers. It has been replaced by advanced techniques

available through HTML5. This has significantly limited its lifespan as a useful and accessible interface. Despite these shortcomings, it is an important early work as it embraced the latest technology of the time and presented the digital collection in ways that would not be possible either in the real world or through search.

Preservation of Favoured Traces

Ben Fry's (2009) *Preservation of Favoured Traces* is a delightfully simple interactive animation that visualises the evolution of additions made to Darwin's text *On the Origin of Species*. The interface displays the actual text of the book in six columns (one for each chapter); as the animation progresses new words are added and others removed, allowing the user to watch how Darwin edited and incorporated more than 40,000 additional words over the six editions (Figure 8). It is fast and easy to understand, relying on colour, a classic data visualisation technique, to represent the various additions. It allows a user to "view [the] text directly, both on macro-level as it animates, or word-by-word as we examine pieces of the text more closely" (Fry, 2009).

As with *ArtScope*, Fry's work relies on the power of a visual overview for an immediate understanding of the data by presenting an effective summary that gives samples of the content whilst also presenting this information in context. This approach relates to Whitelaw's principles of generous interfaces, which will be discussed in Chapter three. *Preservation of Favoured Traces* is an interesting case study in this discussion of collection access techniques, as it is a visualisation of a text—a static dataset—rather than a representation of a cultural heritage collection which could contain 2D or 3D material. However, it shares the same aims by providing access to a rich cultural dataset in a new and engaging manner. Fry came to this project from a creative-code and data visualisation perspective and his background in this field is evident here where the focus is on the data and aesthetic appearance, rather than the importance of the object itself. For example, it is not possible to click on a selected line of text and see the original object—an issue which would not be overlooked on a cultural institution site.

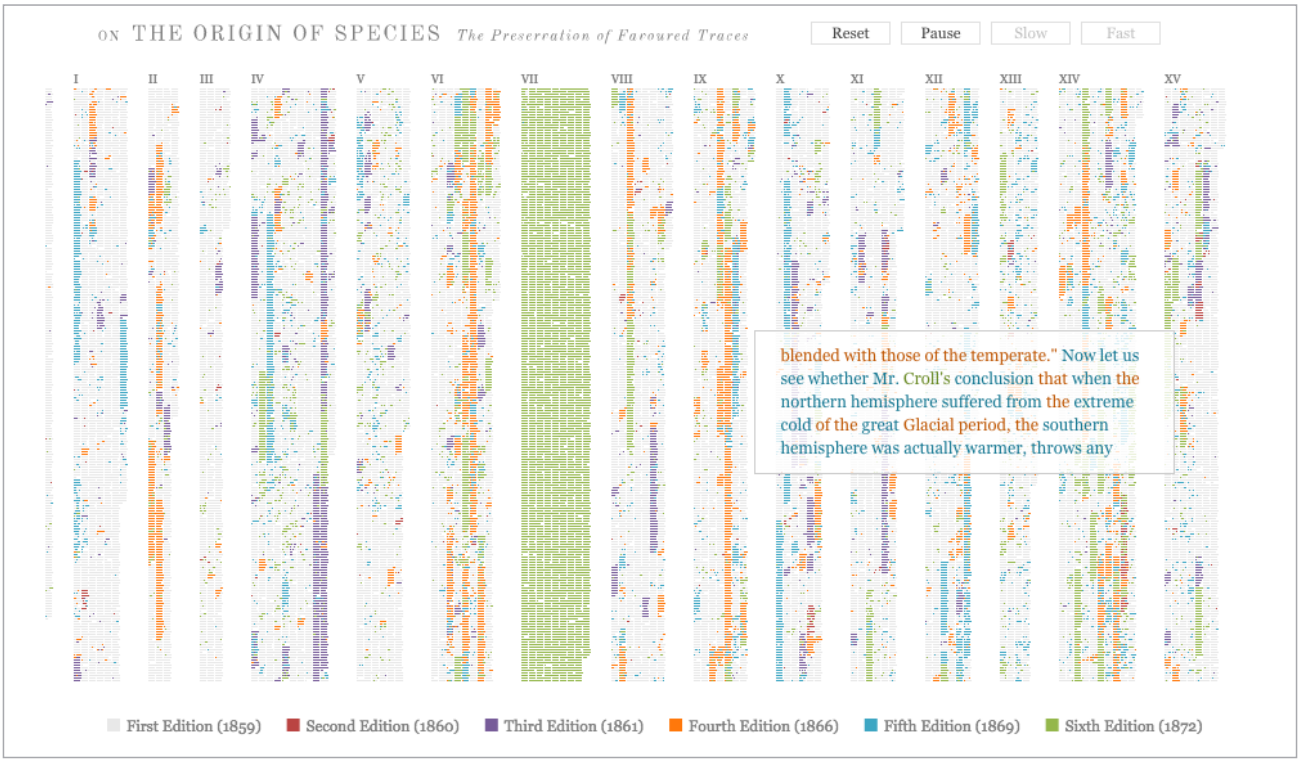


Figure 8
Preservation of Favoured Traces

Preservation of Favoured Traces was developed using Processing[8] java-based language. The original version was available online as a java applet which, as with Flash, required the user to have the browser add-on installed. Recognising how this would limit the lifespan of the animation, it is now available as a HTML5 based animation. By embracing open Web standards, Fry has ensured the work will to continue to function across multiple browsers and devices into the future.

An important aspect of this work is its reliance on a clear and simple aesthetic style. The design is unobtrusive, with only six colours being used to represent the information. The focus is clearly on the content and elements of the interface do not detract from the product. It is a beautiful work and its recent (2015) publication as a poster and a book using the same visual style shows the effectiveness of its aesthetics.

Visible Archive

Visible Archive (Whitelaw, 2009) is a practice-based research project developed by Mitchell Whitelaw for the National Archives of Australia

8 http://processing.org

in 2009. It comprises two components: *Series Browser* and *A1 Explorer*. *Series Browser* presents the user with an overview of all 60,000 series of documents in the National Archive of Australia's collection and *A1 Explorer* visualises the content from just one of those series furthering the visual information-seeking process.

A1 Explorer features a word frequency cloud based on the titles of the data; a histogram of dates and a list of the actual items in that series (Figure 9). When a user hovers over a word in the cloud, lines are drawn to show co-occurrences between terms, thus providing clues about relationships between them. A user can choose to filter the display by focusing on or excluding a term; this then enables them to zoom into the dataset. Focusing on a term will show only records with the given term, whereas excluding it will show all records without that term. This is a powerful technique which allows the user to rapidly move from an overview of 65,000 items to less than 250. They can then receive details on demand and actually see a reproduction of the document they have selected. An interface like this is particularly effective at encouraging exploration as it engages the user through a highly interactive display that reveals relationships that would otherwise remain hidden.

Series Browser is not image based, but instead presents each series of documents as a rectangle with an inner square and an outer band of colour (Figure 10). The inner square is proportional to the number of items in that series and the band of colour represents the amount of shelving it takes up. This is a strong visual pattern, targeting our perception abilities just as Shneiderman had hoped. It invites the user to question the data and investigate the relationships between the different components; a process which will be different for each user and, as a result, lead to greater exploration and understandings of the data (Keim, 2002).

Whilst *Visible Archive* is an important early work, its lifespan is somewhat limited as it runs as a standalone Java applet. This requires the user to download the program and run it locally, rather than visiting it through a Web browser. Technically, it combines both static and live elements. The dataset is a static export of the

Figure 9
Visible Archive: A1 Explorer

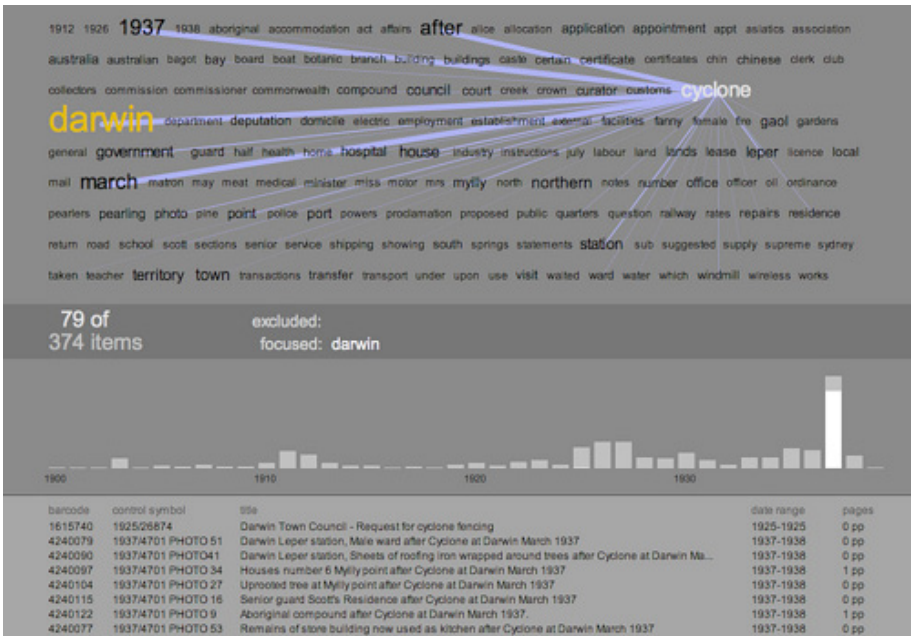


Figure 10
Visible Archive: Series Browser

collection metadata, frozen at the point of time it was created. Whereas the images are loaded straight from the Archives' server. The project was developed by Whitelaw as a research project with support from the National Archives of Australia.

commonsExplorer

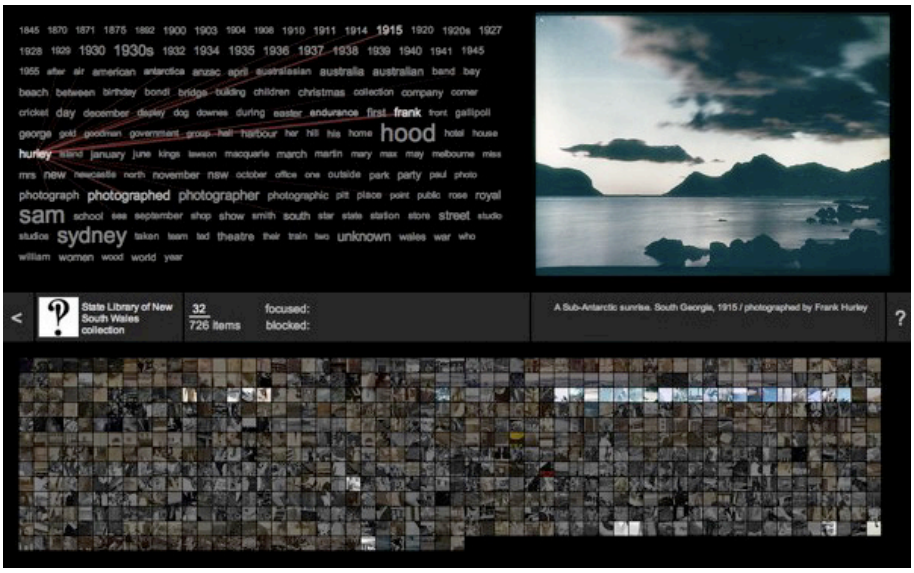
Building upon some of the ideas and techniques advanced in *Visible Archive*, Sam Hinton and Whitelaw released *commonsExplorer* in 2010 (Hinton and Whitelaw, 2010). For this project, the data came from Flickr Commons[9] and, as with *Visible Archive*, it ran as a downloadable Java applet.

commonsExplorer starts with a rich introduction to the data, displaying it in different ways and offering an overview of the collection material (Figure 11). Two key techniques are used here: a word frequency cloud; and a grid of thumbnail images. The word frequency cloud is much like that in *Visible Archive*, and is produced using information from the titles of the various artworks. The font size of the word is relative to occurrence frequency. When you hover over the word, faint blue lines show co-occurrences between different terms; if you click, it will then filter the display. The thumbnail grid is an early attempt at providing an overview of all the images in the collection. Each thumbnail is a perfect square and so the artworks are cropped to the middle to make them fit accordingly. The grid allows the size of the thumbnail to scale up or down according to how many artworks there are in the collection. By cropping the work, a user can see a sample of the collection, which provides contextual clues about that work. Whitelaw explains: "the intention isn't to represent the whole image but to provide some rich but unstructured visual clues: a sort of visual core sample through the whole set." (Whitelaw, 2010). If you select a thumbnail a larger version of the work appears within the current interface, to the right of the word frequency cloud, allowing the user to view the work in context and a further click leads to the original image location on Flickr. There is a clear link between this work and the development of Whitelaw's concept of generous interfaces.

commonsExplorer aims to encourage exploration within cultural heritage data available on Flickr Commons. In keeping with Dörk's

9 Flickr Commons was launched in 2008 and aims to increase access to publicly-held photography collections. https://www.flickr.com/commons

Figure 11
commonsExplorer



concept of the information flaneur there is no predetermined goal, but instead a desire to simply allow free-form exploration to occur.

Walker Art Center

Seb Chan (2011), the Director of Digital & Emerging Media at the Cooper Hewitt Smithsonian Design Museum, described the release of the new Walker Art Center website in 2011 as a "potential paradigm shift for institutional websites." His remark refers to the entire museum website, which was dramatically redesigned to become an online hub that promotes through traffic without trying to trap the user. It is a site that is content-heavy and has parallels with a newspaper website. Alongside this reinterpretation of the traditional museum website came a unique and innovative collection access interface (Walker Art Center, 2011). Upon accessing the page, the user is immediately presented with a screen of high quality images of artworks, with titles underneath (Figure 12). This approach is the epitome of the 'show first, don't ask' principle of generous interfaces. The interface is simple and non-obtrusive, allowing users to focus solely on the artworks themselves. Across the top of the window, the traditional page header content is condensed to one line (with the logo and main site navigation) and below that are a number of options, including the ability to search.

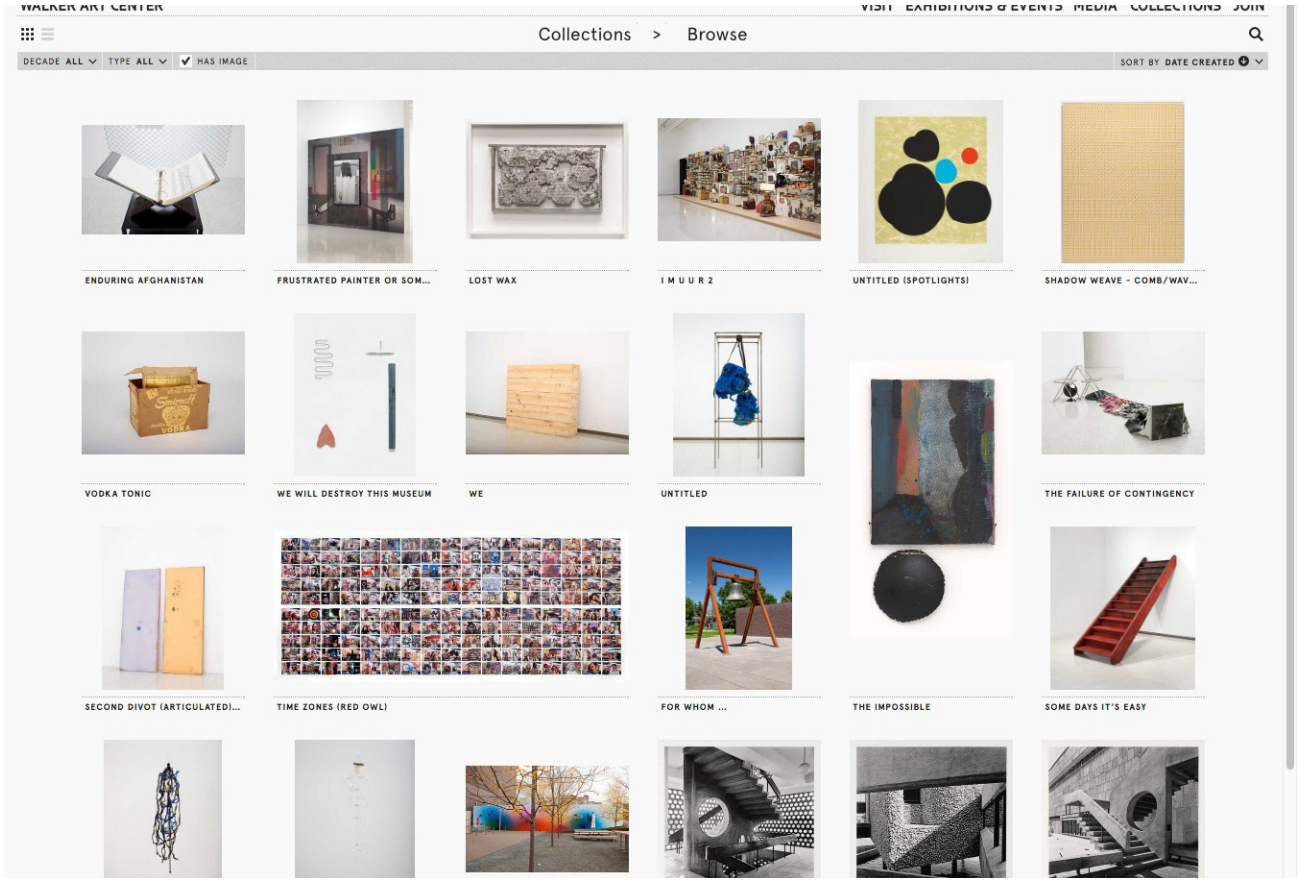


Figure 12
Walker Art Center

By default, the interface displays all artworks, across all decades, sorted by the date they were created. Importantly, the display is already filtered to only show artworks that are imaged.

The pre-selection of the 'has image' box is important because it provides an overview of the collection which is already filtered; in other words, it is not a true representative view of the entire collection. With 'has image' selected, the type filter indicates that there are 4339 artworks, with 'has image' unselected there are 13220 artworks. This results in a dramatically different looking overview.

Images of artworks presented on the main collection pages are organised as a grid, with some artworks breaking out of the standard grid to cover multiple columns and rows. This appears somewhat random. Artworks that are wider than 'normal' always extend across two columns; 'tall' artworks extend across two rows and occasionally some artworks take up both two rows and two columns. This style of display has been strongly influenced by other rich visual Web applications, such as Pinterest, Tumblr

and Facebook. Whilst seemingly insignificant, it is small features such as this, which move the interface away from that of a normal image results page. The interface is responsive and as the width of the browser window is made larger, more columns appear. This responsiveness also applies to images on the individual artworks page, resulting in users being able to see a significantly larger work image than would normally be found.

The Walker Art Centre site embraces the possibilities of new Web technology, built with semantic HTML5 and the latest CSS3 and JavaScript techniques. As with Fry's *Preservation of Favoured Traces*, this ensures that the site will continue to function in any online environment. It is a groundbreaking work, not just because it attempts to redefine the role of the museum website, but also how it achieves this. It is one of the few museum websites that consider the importance of design and the needs of the user when providing access to their digital collections. This is clearly reflected in the design and functionality of the collection's interface.

Google Art Project

The Google Art Project, originally launched in February 2011[10], represents another significant addition to the field (Google, 2011). It allows a user to access data from different cultural institutions in a number of ways: by 'walking' through the actual museum or gallery[11] and focusing on high-resolution images of artworks in situ; by 'Exploring' artworks through predefined categories; and accessing artworks displayed through curated 'Stories' and 'Projects' (Figure 13).

The Google Art Project resurrects the virtual tour of early museum and gallery websites by applying their 'Street View' technology to a new cultural context (Figure 14). Whilst it is an interesting method of providing access to the collection, it is limited temporally, only allowing the user to view what was on display when

10 Google Art Project is now part of the Google Cultural Institute.

11 A return to the 'virtual tour' of early museum and gallery websites, albeit with much improved technology.

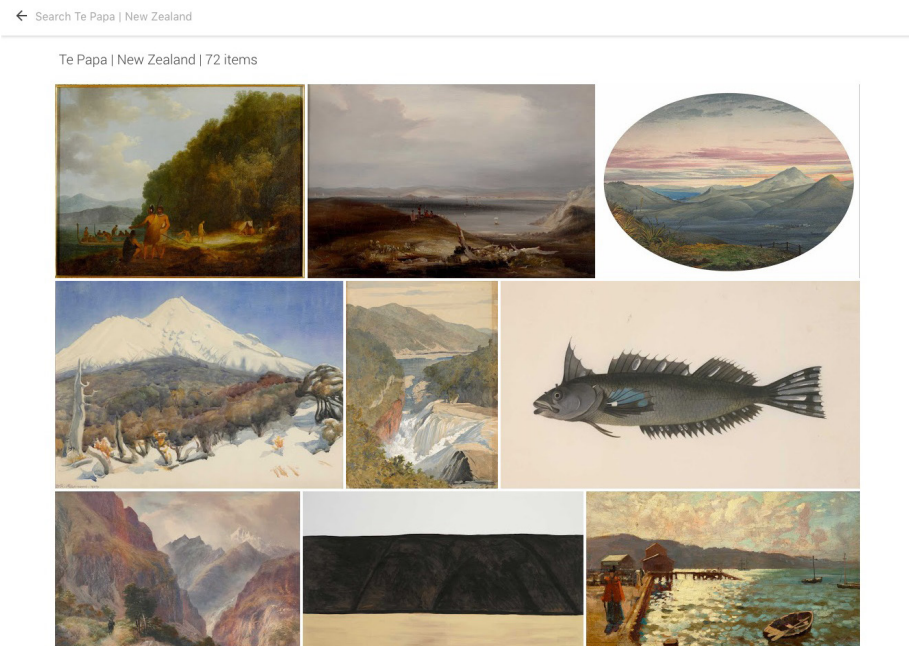


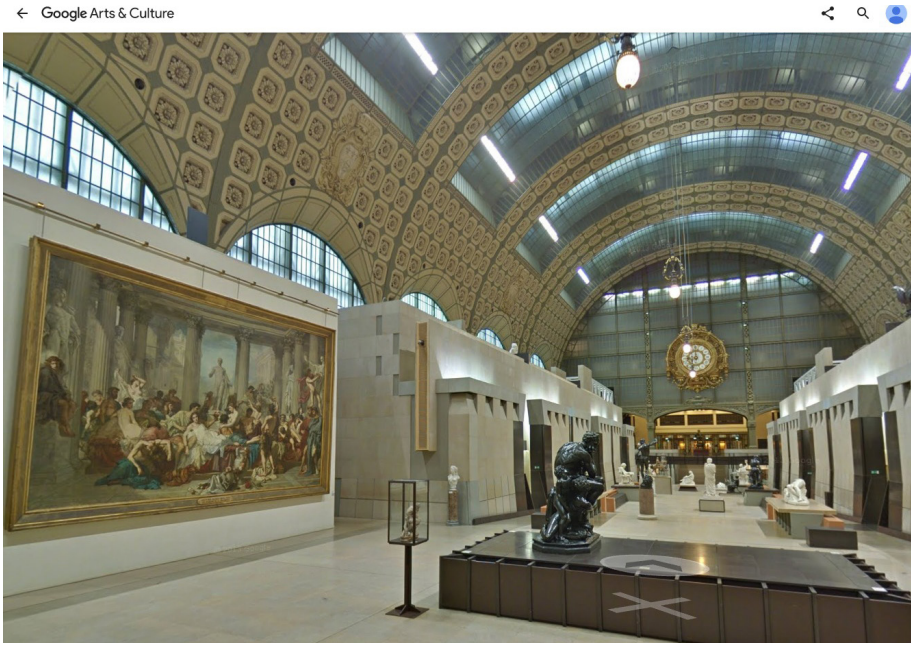
Figure 13
Google Art Project
Te Papa Collection page

it was captured. The 'Explore' functionality groups together artworks from all the partner cultural institutions into a number of predefined categories: artists, mediums, art movements, historical events, historical figures and places. Subcategories are displayed as tightly cropped square thumbnails within a large grid, where is possible to view them alphabetically or through a timeline (in most cases). The diverse range of the subcategories makes the displays visually appealing, considering they are essentially a page of search results.

On the individual category pages, artworks are shown in a grid of randomly sized thumbnails which scrolls horizontally across the full width of the page. The size of the thumbnails varies greatly, from small squares to large rectangles, and the artworks are cropped accordingly to make them fit. In the 'Stories' and 'Projects' sections, the data is presented in a similar fashion. Throughout the Google Art Project interface there is a reliance on displaying content within strong grids with minimal typography and lots of white space. As with the Walker Art Center site, this style is influenced by other visual Web applications and is additionally informed by specifications from Google's Material Design language[12]. It is an engaging way to present artworks and it differs greatly from the established convention.

12 https://material.google.com

Figure 14
Google Art Project
Musee D'Orsay Virtual Tour



The project was launched in Australia in April 2012 with the NGA as the inaugural partner. When this occurred, Peter Naumann, Head of Learning and Access at the NGA explained that it was "an example of the use of new technologies to increase access to art collections" (Naumann, 2012). The perceived benefit for the cultural institution partners is the ability to present items from their collection to a much wider audience, through the dominance of Google, than would otherwise be possible through their own websites.

Whilst there is no doubt that the Google Art Project has the potential to increase access to digital cultural heritage collections through this unique platform, it does have shortcomings. For example, the artworks presented online are chosen by the gallery—providing a glimpse of the collection, or a carefully curated selection, rather than providing an overview of all artworks (only those with copyright clearance could be included). Additionally, the artworks are sometimes presented with only basic contextual information. Later in this chapter, I will discuss how the role of the interface plays an important role in understanding an artwork.

The Google Art Project is one example of an emerging trend within cultural institutions to adopt platforms built by outside groups. The National Library of Australia and many other institutions now upload galleries of images to Flickr, and allow their work to

be ‘pinned’ on Pinterest. These platforms facilitate new ways of interacting with cultural collections, they engage users and encourage exploration. While Pinterest and Flickr don’t allow an overview based approach of the entire dataset, they do embrace the additional steps of Shneiderman’s Visual Information Seeking mantra, including relate and extract, as they allow a user to view related artworks and save them to their own collections, thus personalising them.

Circus Oz: The Living Archive

The Circus Oz: *The Living Archive* (2012) website was developed as part of a research project funded by the Australian Research Council and launched in February 2012. It provides access to a rich multilayered digitised collection of video recordings, photos and other textual resources and embraces many of Whitelaw’s generous interface principles.

Upon accessing the website, the user is given a glimpse into the collection through a number of different boxes displayed on screen (Figure 15). These boxes contain either a single still image, multiple still thumbnails or short repeating videos (rendered as an animated gif), and text overlaid on them. Above the grid of images are four buttons that offer other ways into the collection (for instance, ‘Explore 32 Collections’), and a search box. Hovering over one of these buttons reveals a further contextual clue as to what type of entity each box represents. For example, hovering on ‘View 1436 acts’, adds an orange label ‘Act’ to five of the boxes on screen thus allowing the user to understand that there are multiple acts represented in the overview; conversely this action occurs if you hover over a box, it highlights the other similar entities and the appropriate button to select. This is a useful technique that facilitates a quick understanding of the breadth of the collection. As with the Walker Art Center site, the first page of the interface does not display a full overview of the collection but instead provides a useful glimpse into the wealth of material it holds. It is an engaging style of dynamic display offering multiple ways to access the collection.

By clicking on one of the ‘type’ buttons (also represented

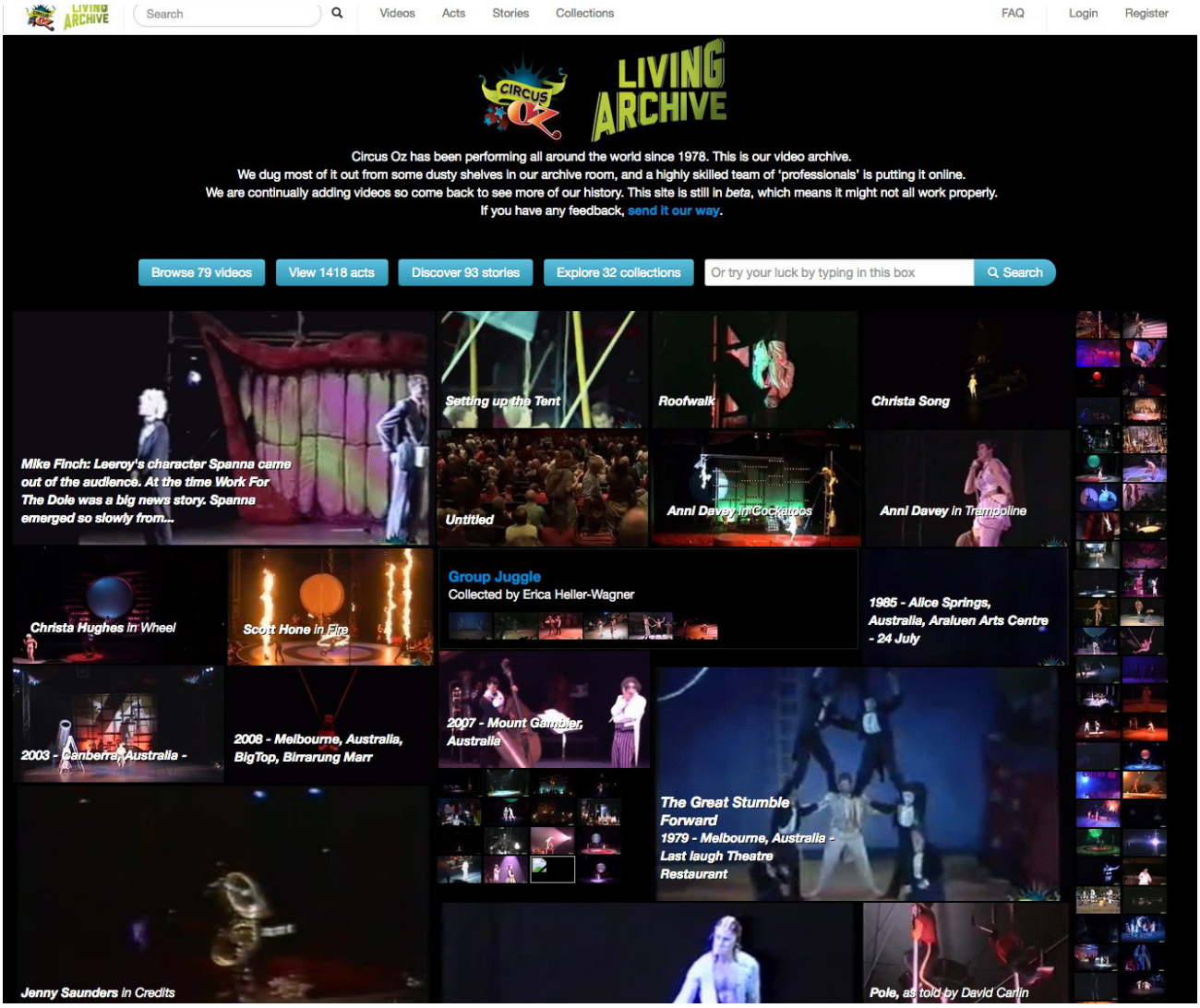


Figure 15
Circus Oz: *The Living Archive*

as navigation items in the header across the top of the page), the user is taken to four different interfaces that each provide a different view of the data and ways to access it. The strongest of these is the 'Acts' page, which presents a dual pane display. Much like *commonsExplorer*, there is a word frequency cloud at the top, and large block of thumbnails below. The opacity level of each thumbnail has been reduced so they appear muted—a common Web design technique employed to encourage a user to hover over the image. As with the display of the similar entities on the front page, hovering over a word in the frequency cloud means the corresponding images below light up (removing the opacity level) and vice-versa. Upon clicking a word, all the highlighted acts remain visible and it is possible to hover over the images to see the name of the act and click to go to its full record. This interface embraces

some of the key aspects of Shneiderman's mantra: 'overview first', 'filter' and 'details-on-demand'. It gives a full overview of all of the acts within the dataset, and encourages the user to interact with the data and see how different types of acts are represented within the collection. Additional interfaces for the other types of acts include a timeline histogram that displays video information, and a three column grid layout layout for 'Stories' and 'Collections'.

Clicking on any individual record, takes the user to an individual record page which has more substantial detail. This page is similar to that of a traditional online catalogue website with entity information and links to other related performers and the like. If a video exists, it is embedded in the page and browsable via segments. The single record page does have some minor improvements: there is the ability to add comments and tags. Clicking on a tag, or an associated record takes the user to an uninspiring search results page. It is clear there is a divide between creating novel visual information seeking interfaces and ensuring the original record can still be accessible via a more traditional display.

Technically, the site is built with the Bootstrap front-end framework, and much of the design style comes directly from there. Bootstrap embraces the latest HTML5, CSS and JavaScript techniques and allows developers to quickly build upon it. Some elements of the design could be refined, for example, by using alternative fonts to Arial and Helvetica, which would alter the user's overall impression and increase the potential for a more positive user experience .

Rijksmuseum

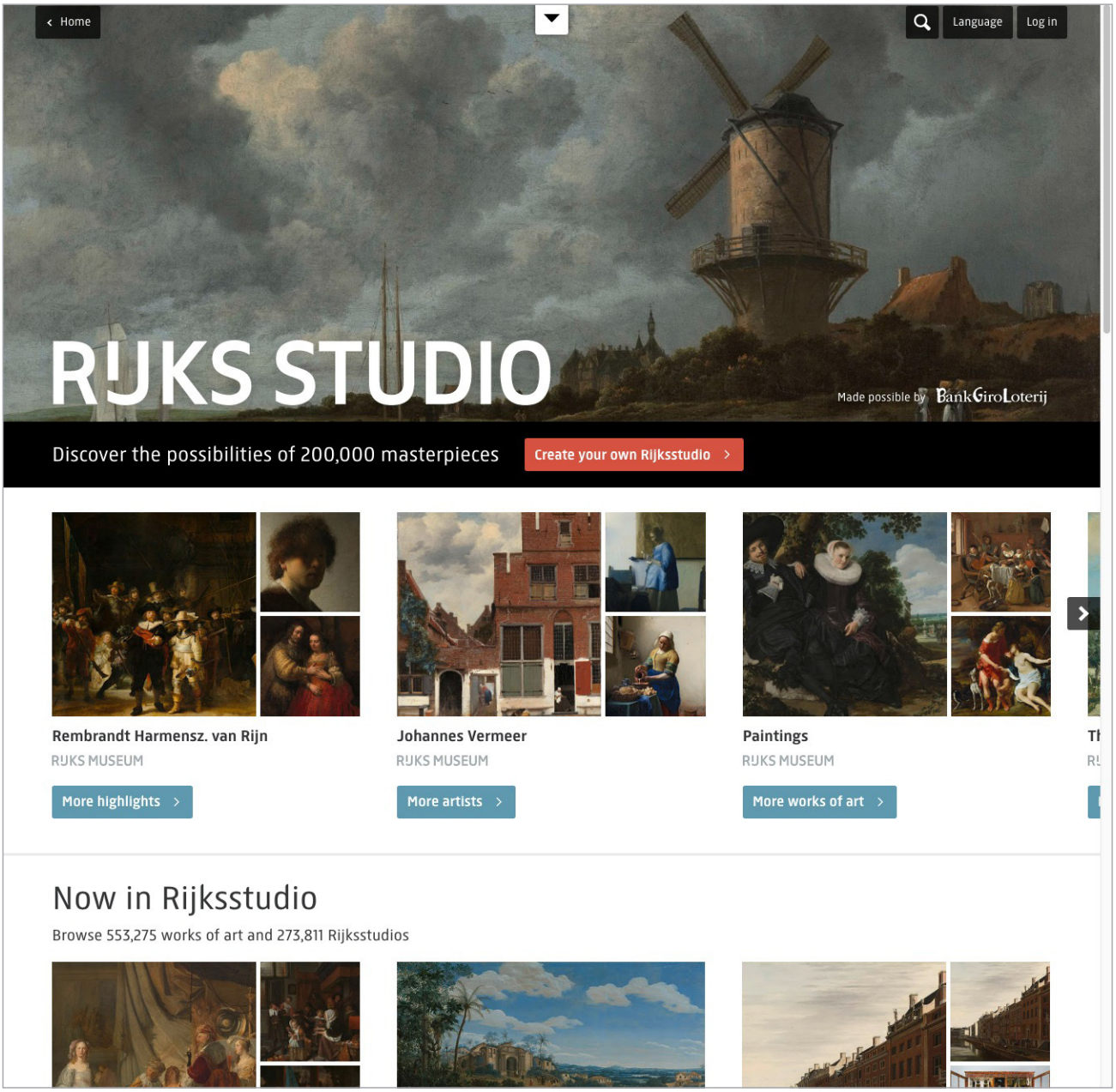
Amsterdam's Rijksmuseum (2012) released their award-winning new website at the end of October 2012, as part of the launch of their refurbished gallery. The site won the *Museums and the Web* 'Best of the Web' award in 2013. When it was launched, it had three different modes which provided access to the collection: the first, the Rijks Studio, "invites members of the public to create their own masterpieces by downloading images of artworks or details of

artworks in the collection and using them in a creative way" (Gorgels, 2013); the second was to 'Explore'; and the third was to 'Search'.

In June 2016, the website was reconfigured and the 'Explore' and 'Search' functionality was moved into the *Rijks Studio* itself (Figure 16). In a blog post discussing the change, they explain how "there is no longer a distinction between collection and *Rijks Studio*" and that publishing user-generated content alongside their own content "is our way of showing that we are OPEN online" (Rijksmuseum, 2016).

There are two key aspects of the Rijksmuseum site that

Figure 16
Rijks Studio



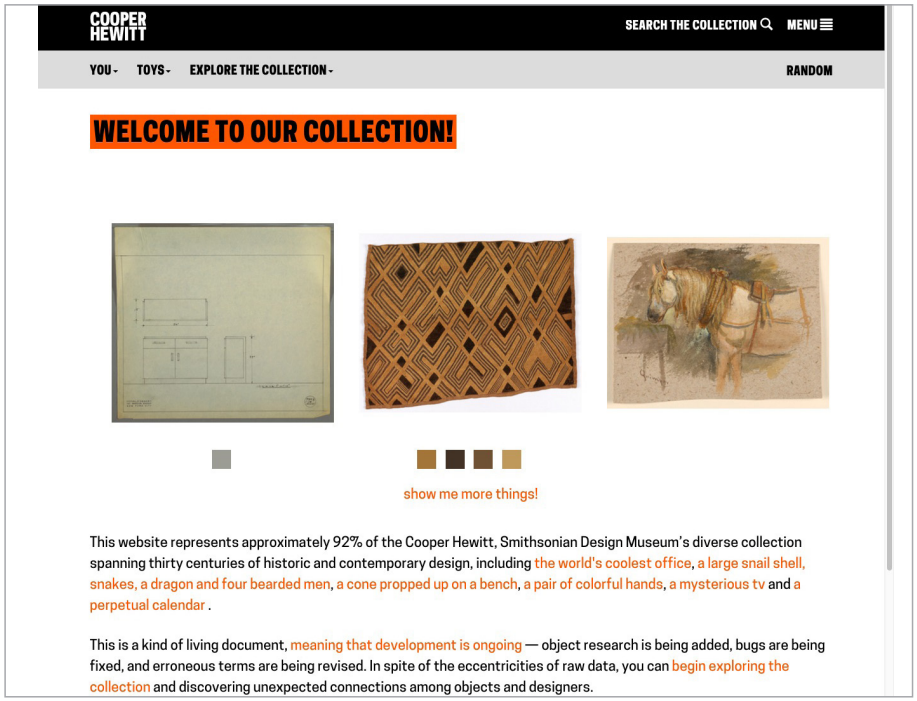
make it stand out in the field. The first, is the open licensing of all Rijksmuseum images. Through the Rijks Studio, users are encouraged to explore, create and share their own collections of works from the museum, thus encouraging deep exploration, with a playful touch. Secondly, substantial attention is paid to the design and user experience. Through the effective layout, careful selection of fonts and a clear colour scheme, the designers have built a site that gives a sense of purpose—it feels fit to modern life and use on various devices. Peter Gorgels, the digital manager at Rijksmuseum, states that it was built with a "tablet first" approach, where the priority is ease of use (Gorgels, 2013).

The new *Rijks Studio* brings a substantial change to the site's functionality, as it promotes exploration and discovery in ways not previously seen. There is now no substantial distinction between content created, and curated, by the Rijksmuseum itself and that created by its users. As with the Walker Art Center, the Rijksmuseum steps away from search as the primary access point and instead encourages the user to choose from a number of thematic entrance points. These are pre-selected, like the 'Stories' in the Google Art Project—Whitelaw (2015) calls them "partial curated" views—and preference imaged artworks only. Whilst they provide a useful starting point they do little to facilitate an overview of the collection.

Cooper Hewitt

The Cooper Hewitt, Smithsonian Design Museum is the preeminent museum and educational authority for the study of design in the United States (Cooper Hewitt, 2014). The current collection interface (Cooper Hewitt, 2016) was first published as an 'alpha' release in 2012 (Chan, 2012), before winning the Best Research / Collections Online honour at the *Museums and the Web* 'Best of the Web' awards in 2013. The review that follows is of the current (2016) version (Figure 17). It must be noted that this discussion aims to highlight key components of the interface which are relevant to my project, rather than providing an in-depth analysis of its many features.

Figure 17
Cooper Hewitt Collections interface



After opening the collections site (https://collection.cooperhewitt.org) the user is immediately presented with a random selection of three items from the collection. Each is represented by a large image of the object and hovering over them shows a 'zoomed in' detailed view of that work, in some cases the dominant colours, extracted from the imaged object, are also shown. Clicking the image leads to the object page, whilst clicking a colour allows the user to see other objects that contain the same shade, presented in a classic search results style display. Below the three images is a playful prompt 'show me more things', a simple strategy, but one that works well in prompting the user to start engaging with collection material.

Throughout the site, the metadata from the objects is used as copy, where it is integrated into a paragraph, rather than just being listed, as would normally be the case (Figure 18). This style of data-driven text is an interesting concept and one I reflect on in Chapter seven.

This technique is deployed particularly well on the individual objects page as it allows the data to tell the story of the object (Figure 19). Additionally, many of these are linked and function as a navigation device, as Sam Brenner, an Interactive Media Developer at Cooper Hewitt, (2015) explains: "by

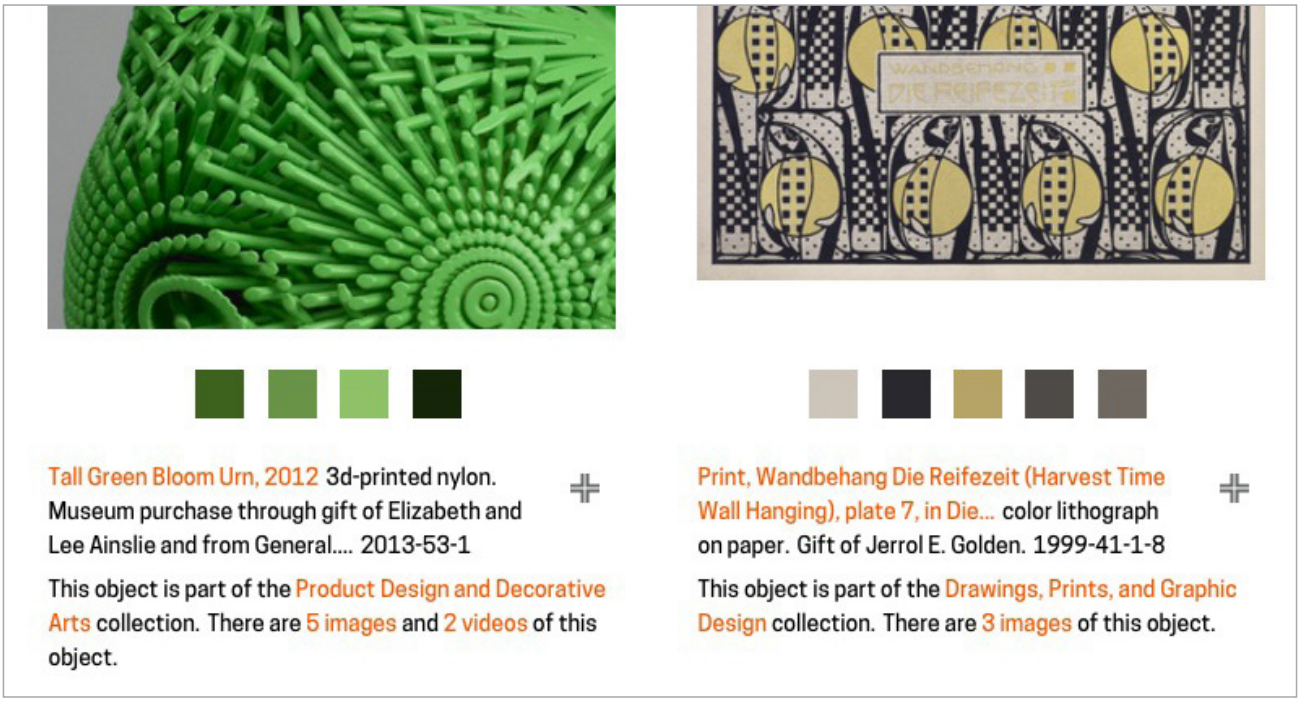


Figure 18
Cooper Hewitt data-driven copy

linking as many relevant first-class items (e.g., objects, people, exhibitions, tags) from an object's individual page, we provided users with an open-ended means of traversing our database that allows users to compose their own narratives."

The leveraging of the metadata is a significant component of the Cooper Hewitt site. By attaching new metadata to existing objects, as demonstrated through the extraction of colours, the user is able to access objects in novel ways. However, it is apparent that the focus of the collections interface is about encouraging exploration through the metadata itself—rather than creating unique access interfaces.

2.8 Summary

In this practical review I have demonstrated that gallery and museum websites still tend to rely on keyword-based search for providing access to their rich collections. In some cases curated highlights are offered as an alternative, however, they do little to provide the user with comprehensive understanding of the collection, nor do they encourage free-form exploration or serendipitous discovery.

There are significant obstacles for galleries and museums to overcome, for example, through funding, negotiating copyright

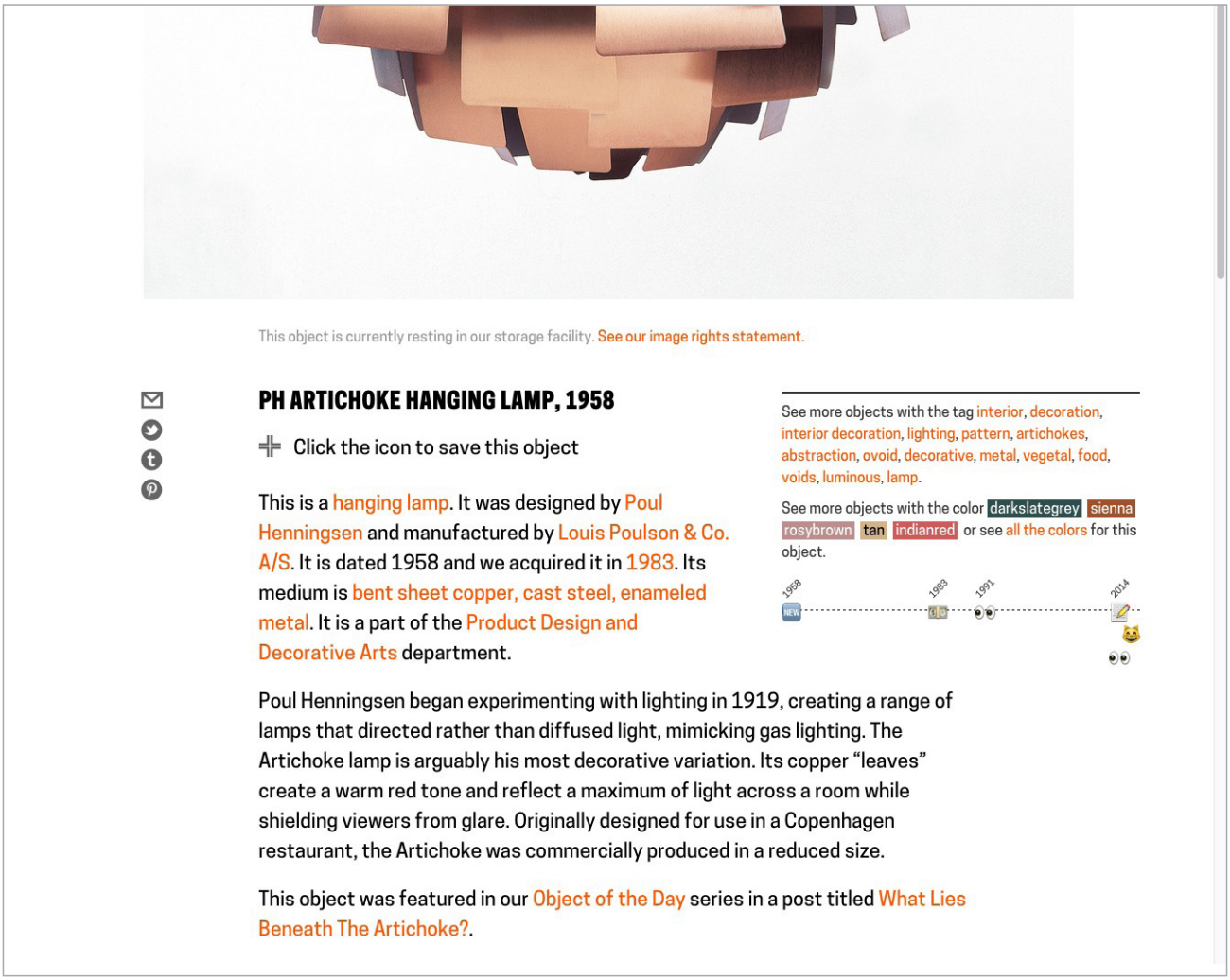


Figure 19
Cooper Hewitt object page

clearance and overcoming a rick-averse culture. However, despite these challenges, I have shown that alternative approaches are emerging. Through the analysis of a number of cases studies I have shown how these tools move away from search and instead start to investigate the possibilities of presenting more engaging methods of collection access. Some cultural institutions, such as Rijksmuseum, Walker Art Center and Cooper Hewitt, have managed to adopt modern Web design conventions and possibilities and developed new collection access techniques in-house, whilst others were created externally for the cultural institutions. In the latter cases, their development clearly reflects the varied backgrounds and interests of their creators as they bring techniques from other fields, such as data visualisation and digital mapping, and apply them to cultural heritage data in order to create novel, exploration focused interfaces.

2.9 Conclusion

In this chapter, I have reviewed the field of practice in which my research project is situated. I have discussed how galleries and museums navigated a path through the emerging technology of the Web by initially creating brochure-style websites, followed by virtual tours, online exhibitions and more substantial online collection access via keyword-based search. I have identified a number of case studies, which I consider to be exemplars from the field, and examined them in detail. Through this analysis, I have shown how they provide new methods for accessing rich cultural heritage collections, by applying techniques from other fields and considering the capabilities of the Web itself. They have informed my own practice by showing the potential of exploration and discovery with digital collections. These tools: *ArtScope*; *Preservation of Favoured Traces*; *Visible Archive*; the Walker Art Center collection interface; Google Art Project; Circus Oz: *The Living Archive*; Rijksmuseum and the Cooper Hewitt collections interface, provide an ample contextual background for my project and show how my own practice-based works are an original contribution to the field.

In the following chapter I will consider a number of theories and concepts from a wide range of literature in order to provide a theoretical framework for the project. In Chapter four, I will describe the research methodology and methods which informed my project, before outlining the creative works in Chapter five. Chapter six focuses on the mixed-method evaluation study and its results. In Chapter seven I will reflect on the practice-based works and address the research questions which governed my project. Finally, in Chapter eight, I will summarise the project and outline its contributions to new knowledge in the fields of cultural heritage and design.

3 Literature Review

3.1 Introduction

In this chapter, I examine literature from a wide range of practitioners and introduce a number of theories which provide the critical basis for the digital interfaces I developed during my candidature.

I begin by introducing the concept of information retrieval (Mooers, 1951) and discuss how it emerged through the need to retrieve information from library systems. These systems required the user to be able to identify an information need (Taylor, 1962), which could then be formulated as a query and used to search a database (Manning et al., 2008). These databases were also used to maintain records in galleries and museums, before being made available online (Butler, 2013; Pierroux, 1998). I will discuss how there are limitations in relying on search, as a user is often unable to specify their query (Belkin et al., 1982) due to a lack of knowledge about what to search for (Donovan, 1997; Eaton and Zhao, 2001; McCrickard and Kehoe, 1997). Following this, I describe the notion of information seeking, and examine how traditional task-based models, such as those by Kirkelas (1983) and Kuhlthau (1991), assume the user has a certain goal. I will argue that these linear modes of information seeking are not reflected in the new ways we seek information and describe how alternative notions, such as exploratory search (Marchionini, 2006), browse (Bates, 2007; Case, 2012), berrypicking (Bates, 1989) and the information flaneur (Dörk et al., 2011), are more appropriate as they embrace our current exploratory-based information seeking experiences. I then turn to discussing visual information seeking, and outline Shneiderman's (1996) Visual Information Seeking mantra, which provides a foundation for the creative work in my project. Building upon this concept I describe "visual information exploration" (Dörk et al., 2012) as a Web-centric model that encourages the combination of visualisation and other information seeking methods. The next part of the chapter considers theories of digital cultural heritage collection interfaces, specifically the role

of the interface, Whitelaw's (2012) generous interfaces, the idea of overview (Greene et al., 2000) and focus+context displays (Card et al., 1999). Finally, I will describe the background of data visualisation and introduce the concept of cultural analytics (Manovich, 2007), before reflecting on the application of data visualisation techniques within the digital humanities (Drucker, 2013).

A note about GLAMs

Wendy Davis and Katherine Howard (2013) describe how the landscape in which our cultural heritage institutions operate has been gradually changing following the introduction of digital technologies. This has led to the formation of GLAM (Galleries, Libraries, Archives and Museums) as an alliance in its own right (Mansfield et al., 2014). David and Howard (2013) suggest "that the practice of digitisation, collaboration, and convergence discussed in GLAM literature might be seen as guiding features or principles of GLAM as a disciplinary field." My work is specifically concerned with researching aspects of digital collections online—especially through art museums and galleries[1], rather than with the much wider scope addressed by the GLAM sector.

3.2 Information Retrieval

The concept of information retrieval (IR) provides a useful starting point for the literature review, as it allows me to introduce the field from which search originally emerged. The term was coined in 1951 by Calvin Mooers and it suggests that information is pre-existent and static, rather than produced during the course of interaction. He defined it as:

> the process or method whereby a prospective user of information is able to convert his need for information into an actual list of citations to documents in storage containing information useful to him. It is the finding or discovery process with respect to stored information. [...] Information retrieval is crucial to documentation and organization of knowledge. (Mooers, 1951)

1 In this Exegesis I refer to museums and galleries in order to provide context for a non-Australian reader.

Manning et al. provide a more concise definition:

> finding material (usually documents) of an unstructured nature (usually text) that satisfies an information need from within large collections (usually stored on computers).(Manning et al., 2008)

IR is a field of study which emerged primarily through the need to retrieve information from library systems. Jones and Willett (1997) explain that early information retrieval (IR) systems were used principally by librarians who had been carrying out bibliographic searches using manual tools such as the card catalogue and universal classification scheme. The earliest computer-based IR systems were built in the late 1940s, and as processor speed and storage capacity increased, they progressed from manual library-based approaches to increasingly automated methods (Sanderson and Croft, 2012). This was due to the creation of new databases which built upon the functionality of the previous card-catalogue systems, albeit with increased computing power (Jones and Willett, 1997).

Whereas IR was previously an activity involving only a few specialists—reference librarians, paralegals, and similar "professional searchers" (Manning et al., 2008)—it is now an activity that everyone has practical experience with, as Web search engines have become the standard method of finding information in modern society (Kowalski, 2010; Manning et al., 2008; Sanderson and Croft, 2012).

Reflecting on Manning's description of IR, a more concise definition could be that it is: the process of finding information from within a database using a search query. Nicholas Belkin (1993) describes this as the "standard" view of IR and argues that it is problematic because of its reliance on two underlying assumptions: first, such a view assumes that the user has a specific information need that can be identified in advance; and second, that searching and selecting texts will fulfil that information need. Belkin argues that these assumptions ignore or devalue the significance of the user's own interaction with the text and support only one form of information-seeking behaviour— search.

In his influential paper *The Process of Asking Questions* (Taylor, 1962), Robert Taylor describes the concept of information needs, which aimed to outline the steps a user completes as they attempt to obtain information. He identified four levels of information need a user passes through before encountering the information system (Bruce, 2005). Taylor (1962) defined them as: visceral, conscious, formalized and compromised. At the visceral level the user recognises a need for information but only in vague terms; the next form is the conscious mental description which could be an ambiguous and rambling statement. At the third level the need is formalized and the user is able to form a properly qualified and rational statement of their question; finally, at the compromised level the question is recast in a manner that the user thinks is acceptable to the search system (Taylor, 1962; Large et al., 1999). The focus at the compromised level is on the user being able to formulate an effective text-based query, with an expectation that they do so whilst understanding the system's limitations. Additionally, Taylor (1962) indicates that a user's information needs may be unconscious or inexpressible, which means that the formation of the search query—the compromised need—will be problematic. This is especially apparent when considering that many users can sometimes only specify what they want once they have found it. This is a view shared by Belkin, Oddy and Brooks (1982) who describe how a user is often unable to specify precisely what their need is, as a request to the system. To overcome this problem, they propose an "anomalous state of knowledge" as an alternative, rather than focusing on an information need.

Within the cultural heritage sector, the influence of the traditional library IR system is clearly evident. Butler (2013) outlines the shift in the way the NGA holds information on artworks in their collections: initially records were held in registration stock books, curator's catalogue cards and worksheets, before being combined into a single database that was made available online. These online databases were traditional IR systems that allowed users to query the system according to their individual interest's and for results to be returned in real time (Pierroux, 1998); this

was not possible when using worksheets or catalogue cards.

Whitelaw (2015) explains that, as the dominant feature in traditional IR systems, search has a deep heritage in the history of computing and is strongly ingrained in the thinking about and practice relating to digital collections. As I have outlined, search is limited because it embodies many assumptions about the user's ability in being able to identify what they are looking for, and to write an appropriate query that will yield relevant information.

Having established an understanding of IR and its benefits and limitations. I will now turn to focus on the role of search in managing access to digital cultural heritage collections.

It is important to not completely dismiss search as an ineffective tool for accessing information. As Kevin Donovan (1997) argues: "search engines are powerful in the hands of those already armed with sufficient information to make them work." However, many academics (Eaton and Zhao, 2001; McCrickard and Kehoe, 1997) agree that for those who do not already have this information they are of little use. David Bearman and Jennifer Trant (1998) explain that the search results should enable users to take the next step in acquiring the information they need, but this is not always the case, particularly if they do not know what they are looking for from the outset. Stefano De Caro, Nicoletta Di Blas and Luigi Spagnolo (2010) determined that one of the central problems with search is that it is only capable of locating a specific piece of information the user can precisely identify. In the domain of visual art collections this situation is demonstrated by Kahn et al. (1997) in their discussion of the Detroit Institute of Art's Visual Resources Art Image Database:

> The search interface is remarkably unforgiving, returning absolutely nothing for an Artist or Title search unless the user types the exact same form used in the catalogue (e.g. only "Demuth, Charles" finds any paintings while "Demuth" or "charles demuth" finds nothing). (Kahn et al., 1997)

It is a catch-22 situation where a user can only search the collection and yield relevant results, if they understand it, and the only way to gain an understanding of the collection is via search. De Caro et al.

further explain how a user who is unfamiliar with a collection will engage in a process of trial and error as they attempt to narrow their results, a process that may quickly lead to becoming lost in the data and feeling disappointed by the experience (De Caro et al, 2010).

The debate surrounding the use of search as an access tool within the cultural heritage sector is not a new one. A selection of different papers delivered at the *Museums and the Web* conferences over the last 20 years highlights this. In 1997, for example, Donovan called for a more engaging approach to online collection access in his paper The Best of Intentions: Public Access, the Web & the Evolution of Museum Automation as he rallied against the "frightful blank search field method of providing access to data". In 2004, Darren Peacock, Derek Ellis and John Doolan found that keyword-based search still remained the dominant method of presenting collection data online, despite it being "totally inadequate as a tool for stimulating knowledge" (Peacock et al, 2004). They explain that the focus of online collections should be on facilitating knowledge rather than just delivering information. At the 2015 conference, Craig MacDonald (2015) argued that it is "possible that a large number of people want to find and view museum objects digitally but have been discouraged from doing so due to poor user experience of existing online-collection interfaces."

Summary

This discussion has shown that traditional IR systems improve upon card catalogue and stock books and can be effective in certain circumstances. However, as they require the user to be able to formulate an effective query, they can impede the possibilities of exploration and discovery within digital cultural heritage collections. This is primarily because, as demonstrated in the practice review, search is often provided as the only way to access the collection.

I have explained that IR is focused on the study of the system and the formulation of effective search queries. It is an approach which aligns with a particular goal-based model of information seeking. However, as I was interested in encouraging

free-form exploration and serendipitous discovery, I needed to consider alternative conceptual information seeking models which are user-centered and allow for more opportunistic and unplanned seeking strategies (Turnbull, 2000).

In the discussion that follows, I will introduce the concept of information seeking before examining a number of task-based and exploration-based models derived from the literature.

3.3 Information Seeking

Information seeking refers to the method used to acquire information, that is, the process in which we purposefully seek information in response to filling a gap in our knowledge (Marchionini, 1997; Case, 2012). As a field of study, information seeking provides a philosophical approach, in comparison to information retrieval which is situated within a technical library-studies based discourse.

In the introduction to his book *Information Seeking in Electronic Environments*, Gary Marchionini (1997) explains how rapid technological development and the importance of information has led to a complex, interconnected and dynamic society that is constantly changing, a point on which many scholars are in agreement (for instance, Case, 2012; Large et al., 1999). The process of seeking information has always been important, and now, with the constant growth of information, it is more pressing than ever before (Large et al., 1999). Information is a valuable resource and acquiring and using information are critical activities that lead to the creation of knowledge (Marchionini, 1997).

Case explains how many traditional models of information seeking, such as those by Kirkelas (1983) and Kuhlthau (1991) describe a linear process, and assume the user has a certain goal with which they will be able to resolve—a view which aligns with Belkin's (1993) view of information retrieval (see section 3.2). However, a linear approach does not allow other non-task-based methods of information seeking to occur. In these cases

we need to consider how alternative models, such as exploratory search (Marchionini, 2006), browse (Bates, 2007; Case, 2012), berrypicking (Bates, 1989) and the information flaneur (Dörk et al., 2011) encourage exploratory modes of information seeking.

Task-based models

In their influential study of literature relating to the information seeking behaviour of people in different professions, Gloria Leckie, Karen Pettigrew and Christian Sylvain (1996) identified substantially different methods of information seeking undertaken by engineers, health care professionals and lawyers. They argue that conceptual frameworks developed over the years have focused on specific professional groups or on a specific aspect of the information seeking process. Case (2012) agrees that there are dozens of information seeking models[2] which focus on many varied fields of professional practice. Some of these models, for example, those by James Kirkelas (1983) and Carol Kuhlthau (1991) seek to describe a task-based process, where the user has a particular goal and follows a fairly rigid path of information seeking until they reach that goal.

Krikelas' model of information seeking was one of the first to explicitly outline a process (Weiler, 2004). The model has four steps: perceiving a need; the search itself; finding the information; and using the information (Krikelas, 1983). Krikelas (1983) argued that "information seeking begins when someone perceives that the current state of knowledge is less than that needed to deal with some issue (or problem). The process ends when that perception no longer exists." It could be considered as a 'problem-based information seeking' model.

Kuhlthau's Information Search Process "depicts a series of cognitive and affective states or behaviours through which people are thought to move as they find and evaluate information" (Case, 2012). The model, created in 1993—before widespread use of the Web—has seven stages: initiation, selection, exploration,

2 Other widely cited models include those by Wilson (1981), Ellis (1989) and Leckie et al. (1996).

formulation, collection, presentation and assessment (Kuhlthau, 1991). Importantly, the model is still aligned with the desire to meet an information need, represented by the user initiating the process.

Krikelas and Kuhlthau's models follow a linear approach, which Jeonghyun Kim (2008) argues is due to information seeking being a "task-based behaviour because the ultimate purpose is a task performance: actions to search are motivated by the overall task, and the degree of satisfaction with the search results will depend on their importance in enabling successful task completion." The task-based approach assumes, as Daniel Backhausen (2012) notes: "the user has a static information need which remains unchanged during the seeking process." It does not allow a user to stray from the path, and there is no ability for freeform exploration and discovery to occur. In their review of information seeking literature, Marian Dörk and his collaborators discovered that 'casual' perspectives, such as those incorporating play and pleasure, are rarely considered in information seeking research (Dörk et al, 2011). Similarly, Whitelaw (2012a) explains how it is an approach that is "increasingly out of step with our ever-more-ubiquitous, casual and everyday experiences of information systems." Clearly, the field of information seeking has limited applications in this investigation.

Exploration-based models

The way in which we seek information has changed dramatically with the advent of the internet, and more recently, with the rise of a multitude of connected devices. We no longer need to visit a museum or gallery to view an artwork; instead we can view information about it from our watches, phones, tablets, laptops and desktop computers. These substantial shifts in how we seek and access information mean there is a need for alternative interfaces that embrace everyday ways of interacting with information. In the next section, I will introduce Marchionini's exploratory search; the concept of browsing; Bates' berrypicking model and Dörk's information flaneur, as four models that are relevant to my work as they enable the conception of new forms of information seeking to emerge.

Exploratory search

Marchionini (2006) describes the concept of exploratory search as an alternative to the traditional task-based information seeking model. He reiterates many of the issues already identified regarding effectiveness of formulating effective search queries, and determines that modern information seeking is a complex and open-ended process, rather than one that is goal oriented. He explains how search interfaces should support exploration by allowing the user to filter the data through the use of facets[3], rather than being forced to enter a query. This is a view supported by White et al. (2006) who argue that search interfaces need to evolve as user requirements change from "using search for lookup to using it to learn, investigate and explore."

Browse

The role of browsing in an information seeking context is pertinent. Browse is often presented as the most common alternative to search, however, from a purely technical perspective, they are the same thing. As Shayam Oberoi and collaborators (2015) explain: "from a technical perspective, just about everything in the online collection is the result of a search: the browse categories return search results for the most common artists or materials." For instance, the function of browsing artworks by the letter H on a digital collection interface will merely run a predetermined search query which returns the results using the same interface as those from a normal search query.

On the other hand, in the literature, Rice et al. (2001) explain how considerable confusion exists between the concepts of searching and browsing, due to the characteristics of browsing behaviour not being well understood. This is because it has come to encompass a wide range of different interpretations across various disciplines and in both technical and theoretical settings. In her paper, *What is browsing—really? A model drawing from behavioural science research*, Bates (2007) examined

3 A recognised modern design pattern. Facets are often available as a way of filtering the search results.

the literature and offered the following definition:

> Browsing is the activity of engaging in a series of glimpses, each of which may or may not lead to closer examination of a (physical or represented) object, which examination may or may not lead to (physical and/or conceptual) acquisition of the object. (Bates, 2007)

This definition aligns with Case's (2012) interpretation that browsing has come to refer to a wide range of different behaviours, from aimless scanning to goal-directed searching; and another view described by Borgman, Hirsh and Walter (1995) that browsing is "an interactive process of skimming over information and selecting choices." These definitions are relevant in the context of my work as they move away from the goal-based approach. They allow a model to emerge where a user is able to scan or glimpse the items in a browsable interface[4] and focus on items as they wish. The user is not completing a specific task, or fulfilling an information need, instead the process of browsing is akin to free-form exploration.

Berrypicking

Marcia Bates' berrypicking model was outlined in 1989 and provides an alternative, nonlinear, approach to information seeking. Bates (1989, 2007) argues that this model is much closer to the real behaviour of information seekers because it is ongoing and transformative. It sees the user employing, and adapting, different strategies as they seek information online. The berrypicking metaphor is apt, as it allows a user to pick information from various locations as needed, rather than following a linear path. The information seeking process is seen to evolve as the user builds their knowledge of the topic and gains a greater understanding of it based on their own interests (Bates, 1989). It is particularly useful when considering that exploratory or serendipitous browsing, are not constant; instead, they emerge as users are provided with the tools to

4 Bates (2007) explains a browseable interface as one that "would consist of rich scenes, full of potential objects of interest, that the eye can take in at once (massively parallel processing), then select items within the scene to give closer attention to.

facilitate these styles of information seeking.

Information Flaneur

Marian Dörk, Sheelagh Carpendale and Carey Williamson introduced another alternative information seeking model in their 2011 paper, *The Information Flaneur: A Fresh Look at Information Seeking*. They describe the information flaneur as a "curious, creative and critical persona", based on the flaneur of Paris in the 1840s famously identified by Baudelaire in his 1863 essay *The Painter of Modern Life*. It too is a useful metaphor that is clearly reflected in our current online experiences, as Dörk writes:

> the flaneur does not methodically navigate the streets,
> nor does he scrutinise everything that crosses his path.
> He is the embodiment of exploration and serendipity...
> he appears to have no goal; rather, experiencing city
> life is his primary aim. (Dörk et al, 2011)

The concept of the information flaneur was devised after the researchers identified a lack of appropriate information seeking models that reflect everyday online information experiences, where the user does not have a specific outcome oriented goal, but instead is responsive to the idea of exploration. For example, a comparable experience in a physical environment is wandering around the library shelves to simply browse the available titles. Dörk and his collaborators are careful to explain that the flaneur does not represent all forms of information seeking but a particular class of practices, goals and motivations that involve exploration and reflection (Dörk et al, 2012). These habits are already reflected in commercial online experiences, for instance, through online shopping, and using social media—where there may be no particular goal—but just to see what there is, to have a look around. It would be wise for museums and galleries to learn from these sectors.

Summary

In this section I have introduced the concept of information seeking and shown how it aims to provide models which explain

the information seeking process in different environments. I have demonstrated how traditional models are task-based and assume the user knows what they are looking for, I have argued that these models do not allow for exploratory, or non-task-based, modes of information seeking to occur. To overcome these issues, I described four models—exploratory search; the concept of browsing; berrypicking and the information flaneur—which enable free-form exploration and serendipitous discovery to occur, by embracing our everyday information seeking experiences.

I will now move to discussing the emergence of visual information seeking and the theories which underpin it.

3.4 Visual information seeking

Visual information seeking is a descriptive label for an area of research that considers the application of information seeking in a visual context (Albertson, 2015). It emerged from the need to provide easier ways for a user to retrieve information from a database. For instance, Ahlberg, Williamson and Shneiderman (1992) created "dynamic queries", which allow a user to formulate a database query with a graphical widget (such as a slider), and see a visualisation of the database and search results. By representing the information graphically, they were able to overcome many of the traditional information retrieval issues previously discussed in this chapter. Building upon their concept of dynamic queries (they renamed them 'dynamic query filters'), Ahlberg and Shneiderman (1994) developed other techniques such as tight coupling (updating the display immediately when a query is entered) and starfield displays (an interactive scatterplot), both of which aimed to support browsing within the interface. They describe how the exploration of large information spaces has remained a challenging task despite advancements in technology, and suggest that their techniques can support exploratory browsing, find patterns and expectations and even make browsing fun (Ahlberg and Shneiderman, 1994).

Ahlberg and Shneiderman's work had a substantial impact

on the emerging field, as evidenced by a paper written by Stacie Hibino and Elke Rundensteiner in 1996, where they describe visual information seeking as a "framework for information exploration where users filter data through direct manipulation of dynamic query filters." The explain how, in their creation of MMVIS (a MultiMedia Visual Information Seeking Environment for Video Analysis), the use of visualisation encourages users to explore the data as well as allowing them to see relationships which were not previously visible by other means.

The success of Shneiderman's work with Ahlberg and Williamson led to the creation of his influential Visual Information Seeking mantra which provides a foundation for the creative work in my project.

Shneiderman's Visual Information Seeking Mantra

At the IEEE Symposium on Visual Languages in 1996, Shneiderman presented the paper *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations* where he described his Visual Information Seeking mantra as a useful starting point for "designing advanced [interactive] graphical user interfaces." The guidelines aim to describe the process users engage in when seeking information in interactive environments and are summarised as: "overview first, zoom and filter, then details-on-demand."

In other words, the overview is the first stage of the process, and it should provide the user with a full representation of the data. Shneiderman acknowledges that it can "be difficult to gain an overview of a collection" but reiterates the need to support it; Zoom allows a user to focus on items of interest, and in some cases may help a user preserve their sense of position and maintain context; at the Filter stage a user is able to remove uninteresting items which allows them to control the contents of the display restricting it to only showing items which interest them; finally, details-on-demand provides the ability to select an item and

receive further information about it. Shneiderman explains how the mantra can be expanded by adding the next natural steps: Relate (view relationships among items); History (keep track of past actions); and Extract (allow the information to be saved, or retained offline). He considers it to be a "mantra" as he repeatedly rediscovered it whilst working on new projects.

When Shneiderman presented this paper in 1996, he was witnessing a change in technology and the rise of the Web. He had visions of an exciting future, where the exploration of information spaces would become commonplace:

> "there is promising evidence that the next generation of digital libraries for structured databases, textual documents, and multimedia will enable convenient exploration of growing information spaces by a wider range of users." (Shneiderman, 1996)

Shneiderman had already produced work with Ahlberg which had started to examine these theories of exploration, and considered how "the older terms of information retrieval... are being pushed aside by newer notions of information gathering, seeking, or visualization and data mining, warehousing, or filtering." He considered that advancements in technology, particularly through "rapid and high resolution colour displays", would enable designers to develop interfaces that target the unique possibilities of the human brain. Shneiderman explains:

> "Humans have remarkable perceptual abilities, that are greatly under-utilized in current designs. Users can scan, recognize, and recall images rapidly, and can detect changes in size, color, shape, movement, or texture. They can point to a single pixel, even in a megapixel display, and can drag one object to another to perform an action. User interfaces have been largely text-oriented, so as visual approaches are explored, appealing new opportunities are emerging. "

He argues that it is time to embrace, and experiment with, visual approaches to interface design. I agree with this view and consider that the development of new visual approaches will bring appealing new opportunities for exploration.

Shneiderman's Visual Information Seeking mantra has been cited over 3600 times according to Google Scholar and is considered to be a "notable theoretical development" in the field of information visualisation (Craft & Cairns, 2005).

Daniel Keim (2002) remarks that the mantra aims to present data in a visual form that allows the user to gain insights, interact with, and draw conclusions from it. Ferreira de Oliveira and Levkowitz (2003) found that interactivity within the visualisation provides mechanisms for handling existing complexity within the data. A user needs to be able to engage with the data, to filter and focus on elements they find interesting; it is this process of engagement and interaction that forms a significant part of the visual information seeking process. Keim (2001) calls it "visual data exploration." The process will be different for every user and leads to greater exploration and understandings of the data (Keim, 2002).

While Shneiderman's work dominates the field it is not without criticism. Craft and Cairns, for example, argue that the mantra has become a prescriptive principle for many information visualisation designers, which is problematic because it has not been subjected to adequate and rigorous assessment (Craft & Cairns, 2005). In a study of 60 papers on the subject, Hornbæk and Hertzum found that while they all refer to the 'overview' as an integral part of a successful visualisation, none explicitly define its meaning (Hornbæk & Hertzum, 2011). Shneiderman uses the word 'overview' in a technical sense in which an 'overview' paints a picture of the whole dataset that the visualisation represents (Craft & Cairns, 2005). For Spence (2011) it is sufficient that a user gain their own overview-based perspective of the data, he claims that it "will not always be acquired rapidly, but may require both time and cognitive effort" (Spence in Hornbæk & Hertzum, 2011).

Visual Information Exploration

Building upon these early ideas about visual information seeking, Dörk et al. (2012) introduce the concept of "Visual Information Exploration" thus providing an alternative model of visual

information seeking to Shneiderman's Visual Information Seeking mantra. It is one that "embraces the Web as both an information space and a visualisation platform" (Dörk et al., 2012). The authors identify how information spaces are quickly changing and argue that new exploration mechanisms are needed in order to make better sense of them (Dörk et al., 2012). In their view combining interactive visualisations with established information seeking methods (such as browsing), will allow a new type of information seeking to emerge. It is a flexible Web centric model that evolves with the development of new techniques and new interfaces. It is a useful theoretical framework for my own exploration-based interfaces as it provides an appropriate context for the emergence of new forms of information seeking, ones that encourage free-form exploration and serendipitous discovery.

Summary

In this section I have introduced visual information seeking and described two concepts which are relevant to my project. The first, Shneiderman's (1996) influential Visual Information Seeking mantra provides a foundation for my creative work. It emerged through earlier work, such as the 'dynamic queries' created by Ahlberg, Williamson and Shneiderman (1992) that allowed the formulation of database queries using graphical elements. The Visual Information Seeking mantra outlines the way in which users seek information in interactive environments and provides a useful starting point for the development of any interactive interface. In addition to this, I have shown how Dörk et al.'s (2012) concept of visual information exploration is a Web-centric model that encourages the combination of visualisation and other information seeking methods, an appropriate model considering an aspect of my research question is to investigate how data visualisation techniques can encourage exploration and discovery.

In the following section I will discuss theories of digital cultural heritage collection interfaces that relate to my project's focus on the AP+P collection data.

3.5 Theories of Digital Cultural Heritage Collection Interfaces

In this section I examine theories of digital cultural heritage interfaces. I do this by introducing the role of the interface in digital environments, before outlining Whitelaw's concept of generous interfaces. I will then consider the role of the overview and describe the importance of Card, Mackinlay and Shneiderman's (1999) focus+context displays.

A note about the interface

In computing, the interface refers to a device or program enabling a user to communicate with a computer (OUP, 2010). In the context of my work, I refer to the interface as the elements within the browser window the user can see and interact with, that is: the actual HTML, CSS and JavaScript that generates the page.

The Interface

The importance of the interface in our digital environment cannot be overestimated. Whitelaw (2012a) argues that interfaces matter now more than ever, because increasingly a user's primary experience of a collection (of artworks for instance) is in its digital form. Butler (2013) agrees with this point and observes how a generational change has led to online representations being the major source of information. Oomen, Baltussen and van Erp (2012) have remarked that just pointing to a database with records no longer suffices. Instead the interface plays a crucial role, not just as an access tool, but also as a way of supporting how we see and understand the artwork itself. James Davies explains as follows:

> Online the interface in effect plays a similar role to the frame, the glass, the label, the map, the wall and so in in the gallery. These can either support or distract from an artwork, and many of our existing collection websites do not support the display of artwork very well because we only consider these digital reproductions as mere references to the real thing. (Davies in Proctor, 2011).

When viewed online, each artwork becomes a self-contained object that can only be interpreted by the information that accompanies it. We need to step away from simply replicating a physical artwork online, and instead rethink these displays and the possibilities provided by the new digital medium (Cunliffe, 2001).

Issues surrounding the display of artworks online have a historical basis. As Joanna Sassoon suggested in 2006, when digitising material in the past, the focus was on content, not context (Sassoon, 2006). Nicholas Serota, former Director of the Tate Museums, explains the reason behind this approach: "the first generation of museum ... [websites] were concerned with quantity of information and getting as many objects online as possible" (Serota in Proctor, 2011). However, as Donovan (1997) remarks, "people do not care so much if you have a certain painting... they are fascinated by all the context, the history, associated with it." Serota suggests that a second generation of online collection websites is now emerging and their focus is "on depth and quality of content" (Serota in Proctor, 2011). Therefore if the experience a user has with a collection is primarily online, it has become necessary to develop interfaces that support the display of the artworks whilst maintaining context and creating high quality connections. This view is supported by Grady (2006) who explains that "visual displays are important not just because they confirm or disconfirm testable theories, but because they also generate knowledge that would be unavailable in any other way."

Generous Interfaces

At the 2012 International Congress of Archives Conference in Brisbane, Whitelaw presented a paper entitled *Towards Generous Interfaces for Archival Collections* in which he outlined, for the first time, a conceptual framework he characterised as "generous interfaces." He proposed that the ethos of 'generosity' can provide a "critical perspective on current and future collection interfaces" and argued for it to be be considered as a guiding principle when designing digital collection interfaces. Whitelaw (2012a) explains:

> Our digital collections are certainly large, abundant and ample; and the charters of our cultural institutions place a high value on sharing these riches liberally with the public. Generosity seems to be very much in line with the aims of our cultural collections. (Whitelaw, 2012a)

The concept of generous interfaces emerged through the production of a number of creative works by Whitelaw, including *Visible Archive* (2009), *Manly Images* (2012b) and the first three interfaces in my research project (*Works and Networks*, *Decade Summary* and *All Artists*—which were developed in collaboration with Whitelaw). The principles it outlines: show first, don't ask; provide rich overviews; provide samples; provide context and share high quality primary content, were directly informed through the production of these interfaces.

In the practice review I illustrated how, when it comes to online representations of digital collections, the idea of generosity is almost never present. A generous interface can offer an alternative approach to collection access, where it may simply encourage free-form exploration or it may provide an overview of the digitised collection. Either way, the aim is to volunteer the information to the user, rather than forcing them to seek it.

As mentioned previously, Whitelaw (2012a) describes a number of points to consider when creating generous interfaces; a brief summary of these follows and they are illustrated in more detail in Chapter six.

Principles of Generous Interfaces

Show first, don't ask. Rather than forcing the user to enter a query, volunteer rich information to them.

Provide rich overviews. Shneiderman's Visual Information Seeking mantra sees the overview as the foundation of an exploratory visualisation. A generous interface should give the user a rich overview that allows them to orient themselves in relation to the collection.

Provide samples. It is hard to display an entire digital collection, however samples of the primary content can be shown to give contextual clues and encourage exploration.

Provide context. We should always show and maintain context for the item by displaying the structure and relationships within the collection.

Share high quality primary content. We should always provide easy access to the original high quality primary content, be it a larger image or the full catalogue reference in the standard collection interface.

Overview

The overview has emerged as a key concept throughout the literature. Shneiderman sees it as the starting point of any information seeking process, the aim being to allow the user to gain an immediate summary of the whole collection (Shneiderman, 1996). Stephen Greene, Gary Marchionini, Catherine Plaisant and Ben Shneiderman (2000) agree and explain how "an effective overview provides users with an immediate appreciation for the size and extent of the collection of objects the overview represents." Keller and Tergan (2005) argue that a visualisation starting with an overview can be particularly effective because the user immediately appreciates the size and diversity of a collection (Hornbæk & Hertzum, 2011), whilst, importantly, revealing what kind of objects are not in the collection (Greene et al., 2000).

However, taking Stamen's "show everything" (Jones, 2009) approach, as demonstrated in *ArtScope*, can be overwhelming and limitations in the data, and the sheer size of the collection, can affect the strength of the representation (Greene et al., 2000). This results in the construction of an overview becoming a compelling design challenge in itself. I will reflect on these challenges in Chapter six.

As outlined in the practice review, many museum and gallery websites do not support exploration of their collection. In their paper Previews and Overviews in Digital Libraries: Designing Surrogates

to Support Visual Information Seeking, Greene and co. (2000) argue that "interfaces for digital resources regularly fail to provide honest representation[s] of what they include", they suggest that many users do not, and cannot, understand the structure and type of materials available in an online collection, because there is a lack of interfaces that provide appropriate access. The development of interfaces that include or combine previews—a more focused representation of the collection—and overviews, can help to overcome these issues. Greene et al. (2000) describe how an overview can consist of previews, and a preview might "serve as [an] overview for more fine-grained objects". The aim of the techniques is to provide users with methods to discover objects of interest and better understand the scope and nature of the digital collection. Greene and his collaborators' introduce the concept of the information surrogate, which will incorporate aspects of an overview or preview. Surrogates are compact and browsable abstractions of the primary content and draw upon library and information studies (Whitelaw, 2015). Greene et al. (2000) explain how an overview is constructed from, and represents, a collection of objects of interest; whilst a preview is extracted from, and acts as a surrogate for a single object of interest. These extracts could be literal, such as an artwork thumbnail; or original creations that convey meaning from the primary object. In other words, an overview provides a focused view of a digital collection, and a preview can provide context of the objects themselves.

Greene and co. (2000) outline a number of guidelines which I have found helpful when applying the concepts of overviews and previews in novel contexts. They include:

> using multiple surrogates to support different user styles and experiences
>
> using multiple levels of surrogates and displaying them on a single screen to show hierarchy
>
> using surrogates to inform users about size, extent and availability of collections or objects
>
> leveraging data types—using visual surrogates for visual data
>
> (adapted from Greene et al., 2000)

In addition to these guidelines, Dörk and his collaborators also provide suggestions on how to design interfaces that allow and encourage exploratory behaviour to occur. These align closely with the concepts outlined by Greene et al. For example, Dörk describes horizontal exploration, which involves the user exploring the information from a high level (akin to a birds eye view), through an overview providing orientation; and vertical immersion, where a user follows their interests and becomes immersed in a more detailed level of the data. There are clear similarities between horizontal exploration and the overview concept, and vertical immersion with the concept of the preview, they are all techniques that can help to fulfil the needs of the modern information seeker (Dörk et al, 2011).

Finally, as Kahn, Lenk and Kasman (1997) have written,

> "an overview or other representation of these electronic collections is a useful orientation and navigation tool. Without it, the viewer is at a tremendous disadvantage."

It is clear that the overview is an important part of a digital interface which seeks to encourage exploration and discovery.

Focus+Context

In addition to the role of the overview in the interface, I have been particularly concerned with the development of focus+context displays. The second and third stage of Shneiderman's visual information seeking mantra states 'zoom and filter, details on demand'. If we consider how these stages are applied in a traditional catalogue-based collection access interface, it is evident that focusing on an artwork (from the search results) directs the user to a new page providing the details on demand. However, by going to a new page, the user now loses the context of the previous view. The focus+context approach is an information visualisation technique which was developed to help overcome this problem. Card et al. (1999) describe this technique as one where the user is simultaneously provided with both the overview (context) and detailed information (focus). Cohen and Brodlie (2004) explain how the challenge is to "find a way of looking at a high level of detail at this area of focus,

without losing the overall context." Much of the literature (Lamping et al., 1995; Rao and Card, 1994; Card et al., 1999) refers to the creation of focus+context techniques in software that creates information visualisations, although, in my work I investigate how to create digital interfaces which combine focus+context displays with Web-based interfaces, rather than just focusing on visualisation techniques.

These ideas about generosity, overview and the importance of context parallel the methods some galleries and museums use to display their physical collections. For example, the Victoria & Albert Museum embraces the idea of an overview in their ceramics collection by providing a 'visible storage display' (Victoria & Albert Museum, 2011), filled with ceramics from their collection (Figure 20). A visitor is able to see how various techniques have developed and changed over time as well as identify connections between different artists and styles. This style of display allows the visitor to gain both a full overview of the collection and the many varied objects within it. If such a technique can be so fulfilling in a physical context, it seems unfortunate that the same levels of generosity are not replicated online.

Free-form exploration and serendipitous discovery

In the context of this discussion about exploration and discovery, it is important to introduce and define the concept of 'free-form exploration and serendipitous discovery'.

A definition of 'free-form' states that it does not conform to a regular or formal structure or path (OUP, 2010), whilst serendipity is defined as "the process of making fortunate discoveries for which someone was not looking" (Taramigkou et al., 2003).

When I refer to the phrase 'free-form exploration and serendipitous discovery', I am specifically concerned with promoting exploration within digital interfaces. As such, the concept can be considered as a process of interaction where the user is not following a linear path and has no specific goal. Instead, they are

Figure 20

V&A ceramics display

simply exploring the data and potentially finding unexpected objects of interest. It is the chance discovery of these items that triggers feelings of surprise and delight, or, in other words, a serendipitous discovery. I suggest that in the context of this research project, these feelings are central to the notion of serendipitous discovery. Chan (2007) and Taramigkou et al. (2003) agree that chance encounters can lead to advancements in knowledge, whilst Hangal, Nagpal and Lam (2012) describe serendipitous discoveries as being important because they entertain and captivate users. Chan (2007) explains how the design of an environment, be it physical or digital, can be made to encourage serendipity: "it is increasingly important to offer visitors opportunity for serendipity [...] to retain their attention and to encourage them to explore the Web site."

It is clear that if the design of the interface can provide users with the opportunity to engage in free-form exploration, then serendipitous discoveries are likely to eventuate. I will demonstrate how this occurred when discussing the results of the evaluation study in Chapter six.

Summary

This section discusses relevant theories of digital cultural heritage interfaces. I have described the role of the interface and identified how, in an online environment, it is a significant factor in determining how a user is able to understand and interpret an artwork. Through the introduction of Whitelaw's notion of generous interfaces, I have outlined additional principles to consider when creating exploratory-based collection interfaces. I refer to the role of the overview as an important aspect within any interface which aims to encourage exploration. Effective overviews can serve multiple purposes as they allow users to orient themselves within a collection and can be a useful navigation tool. Finally, I describe how focus+context displays embrace aspects of Shneiderman's Visual Information Seeking mantra, by encouraging the creation of displays where both the overview (focus) and detailed information (context) are displayed at once.

In the final section of this chapter I examine literature surrounding data visualisation and its role as a contemporary field of cultural practice.

3.6 Visualising Culture

The visual representation of data is not a new concept. Data visualisation has its roots in cartography, scientific diagrams, charts and graphs (Friendly and Denis, 2001). A map, for example, is a visual representation of the geographic location of land and sea. The aim of a visualisation is to present data graphically and leverage the cognitive and perceptual capabilities of the human brain (Vande Moere, 2007) in order to gain insight. Visual components, such as: size, colour, shape and scale can be used to help reveal patterns and relationships within datasets, that might otherwise go unnoticed if the same information were represented textually (Dörk, 2012). Andrew Vande-Moere considers data visualisation works to have a single purpose: to increase the human understanding of data (2007); as a result, it is a varied form of cultural practice involving a wide

range of influential researchers and designers.

Data visualisation techniques are used across many different fields, but are perhaps most prevalent when applied to scientific or statistical data (Friendly and Denis, 2001). By 2007, researchers had identified a shift away from these traditional scientific visualisations towards visualisations of abstract data (e.g. Vande-Moere, 2007; Manovich, 2007; Rodenbeck, 2010). Manovich (2007) argues that if scientists and government agencies can use data visualisation as a new way to generate knowledge then it must be possible to apply the same approach to understanding cultural data; he sees visualisation as a way of "generating new approaches for studying cultural history." He coined the phrase 'cultural analytics' to describe

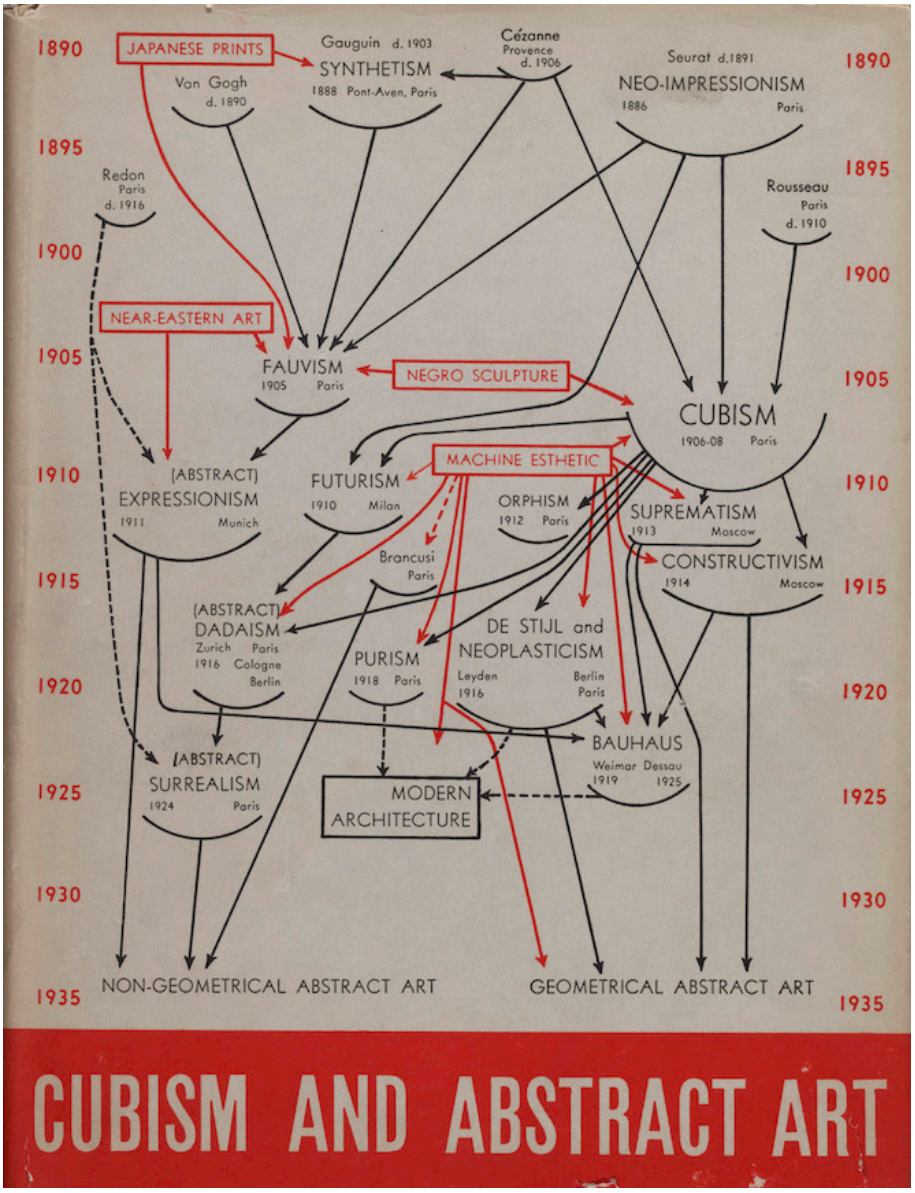


Figure 21
Barr's cubism and abstract art chart

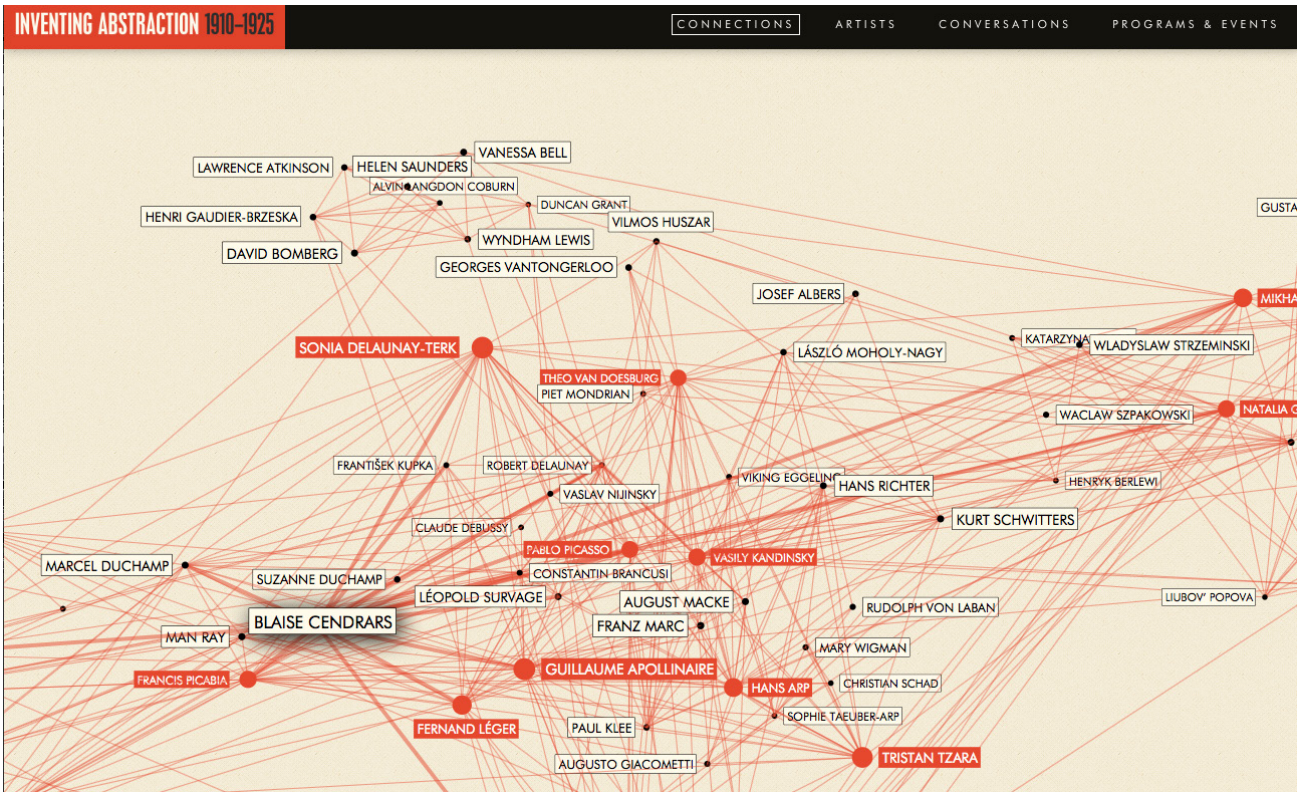


Figure 22
MOMA's *Inventing Abstraction* interactive visualisation

this new discipline dedicated to the analysis and visualisation of large cultural data sets (Manovich, 2007). It is evident that Manovich's argument has been validated, with Jefferson Bailey and Lily Pregill writing in 2014 that "applying digital visualization techniques to cultural heritage data sets is celebrated as a new and innovative research methodology" (Bailey and Pregill, 2014).

An early data visualisation work using cultural heritage data was produced by Alfred H. Barr, Jr. in 1936. The work, which takes the form of a flow chart, was produced for the exhibition *Cubism and Abstract Art* held at the Museum of Modern Art, where Barr was director (Bailey and Pregill, 2014). It uses a number of different visualisation techniques (for instance, size, colour and connected paths) to illustrate the evolution of modern art over the previous 45 years (Figure 21). Tufte explains how it "simultaneously served as a beautiful cover for the catalogue, a table of contents for the show, an organizing history of the art displayed in the museum, and a symbol of the entire enterprise" (Tufte, 2006). From Tufte's remarks it is obvious that an effective visualisation can serve many purposes, regardless of the context in which it is used.

While Barr's chart is a static visualisation, it aligns with some of the ideas already discussed in this chapter. For example, it provides an immediate overview and highlights the relationships that occur between the different artists and art forms. In 2012, the chart was re-created, and expanded upon, as a modern interactive data visualisation for the exhibition *Inventing Abstraction*[5] at MOMA (Figure 22).

Within the data visualisation community there has been considerable debate about the function of a data visualisation work (Few, 2011; Bertini, 2011). Two main approaches have emerged: exploratory data visualisations present the data 'as is' and encourage the viewer to build their own understandings of the data; whereas in communicative visualisations[6] the goal of the designer is to present a particular view, tell a specific story, or to use the data visualisation as a persuasive medium (Ennis Butler et al., 2011). Stephen Few (2012) calls the latter approach "data art", and argues that when we fail to distinguish the two, it can cause confusion and harm.

Florian Kräutli (2016) describes data visualisation as one of the essential research methods for the digital humanities. However, he notes that established paradigms for visualisations in the sciences do not necessarily translate to humanities data, and may need to be reconsidered. This view is supported by Johanna Drucker (2011) who writes that "the sheer power of the graphical display of information visualisation… seems to have produced a momentary blindness among practitioners who would never tolerate such literal assumptions in textual work." Drucker is troubled by the direct adoption of methods from these fields and critiques their implementation in the digital humanities community. She argues how, in work by Shneiderman and others, principles for design and display are "premised on a pragmatic, but highly mechanistic, analysis of a user's abilities to process information effectively" (Drucker, 2013). These values do not align with those from the humanities and instead of relying upon them,

5 http://www.moma.org/interactives/exhibitions/2012/inventingabstraction/
6 Also referred to as infographics

Drucker argues that the digital humanities community should reinvent interface design in a form "suited to its critical principles" (Drucker, 2013; Whitelaw, 2015). These issues stem from a problem outlined by Fiona Cameron and Sarah Kenderdine (2012):

> In the last few years much of the discourse about the relation between cultural heritage and digital technology has been descriptive and introspective, focusing on projects and their technical considerations. … These fora continue to make valuable contributions to the ongoing development of the sector, although they are yet to foster a body of sustained critical thinking about the meanings and implications of the apparent transformations, challenges, and possibilities posed by communications technologies. (Cameron and Kenderdine, 2012)

Similarly, Kräutli (2016) suggests that this apparent lack of rigour in the digital humanities can also stem from the researcher's unfamiliarity with visualisation tools and a lack of critical understanding of them. He describes the need for visualisation tools and methods which fit the characteristics of humanities data and research, but claims that it is also necessary for the researcher to become aware of the mechanisms of data visualisation and how they shape humanities knowledge production.

Summary

This section has introduced data visualisation and discussed its emergence as a field of contemporary cultural practice. I have explained that it has moved beyond its roots in the sciences, and been embraced by academics and researchers as a new way of generating knowledge within abstract datasets. I align my research practice with the development of exploratory data visualisations and emphasise the role of the user in being able to build their own understandings of the data. Finally, I reflect on the application of data visualisation techniques within the digital humanities and identify the need to treat these techniques with scepticism due to a perceived lack of understanding about them.

3.7 Conclusion

In this chapter I have introduced and analysed a wide range of literature which provides a critical basis for my research project. I began by describing how information retrieval systems relied on the user being able to identify an information need and formulate an appropriate query to retrieve information from a database system. I have outlined how these databases were used to maintain records within galleries and museums before they were placed online, leading to the dominance of search. I have described the limitations of search as an access method, as it assumes the user knows what they are looking for, despite there being significant research suggesting otherwise.

I have examined how task-based information seeking models leave little room for exploratory modes of information seeking and have introduced a number of alternative models, such as exploratory search (Marchionini, 2007), browse (Bates, 2007; Case, 2012), berrypicking (Bates, 1989) and the information flaneur (Dörk et al., 2011), which enable non-task-based information seeking to occur. I have shown how concepts of visual information seeking, particularly Shneiderman's (1996) Visual Information Seeking mantra—which provides a foundation for my work, and Dörk et al.'s (2011) Visual Information Exploration outline techniques to consider in the development of exploratory interfaces. I have analysed how theories of digital cultural heritage collection interfaces build upon these ideas about visual information seeking by describing the role of the interface, introducing the notion of generous interfaces (Whitelaw, 2012), the importance of the overview and how focus+context displays can combine many of these techniques. Finally, I examined how data visualisation has become a field of contemporary cultural practice and described some of the issues surrounding its application in the context of digital humanities.

Through this literature review I have demonstrated how my research project is situated within a rich theoretical framework, encompassing theories and concepts from many different fields of practice. In the next chapter I will introduce the research

methodology and methods which informed my project. In Chapter five I describe the creative works, before outlining the evaluation component of the project in Chapter six. Chapter seven provides reflections on the practice-based work, through a discussion that addresses my research questions. In Chapter eight, I summarise the project and outline its contributions to knowledge.

4 Methodology

4.1 Overview

In the previous chapters I discussed the theoretical and practical works that provide a formal context for the interfaces I created. This chapter will introduce practice-based research as the methodology that underpins my research project. I will then consider how the adoption of a number of different methods allowed me to address my research questions effectively. I will refer to: the braid, sketching in code, primary and secondary research, prototyping tools and mixed-method evaluation. I will conclude this chapter by presenting a readaptation of the braid to illustrate how it encapsulates the combination of methodology and methods I have employed in the project.

4.2 Research Methodology

Throughout my PhD I have employed a practice-based research approach where, as Linda Candy (2006) explains, "an original investigation [is] undertaken in order to gain new knowledge partly by means of practice and the outcomes of that practice." Candy (2006) describes how a practice-based research degree specifically includes an artefact, or the record of an artefact as an integral part of their submission; a view supported by Rust et al. (2007) who write that "issues, concerns and interests can be examined and brought out by the production of an artefact." In other words, knowledge is produced through both the creation of the artefact (practical works) and from the artefact themselves. This approach is relevant to my research project for two key reasons: firstly, my background is as a practitioner; and secondly, I wanted to explore how the form of practical work could contribute to new understandings of the collection content itself—gained from using the interfaces.

I come to this research project with a professional background in Web design and development. Whilst undertaking this project

I have worked concurrently as a lecturer in Media Arts and Graphic Design and as a freelance Web developer. I was brought up within the visual arts community by parents who, at times, were both employed by the National Gallery of Australia; as a child I frequently travelled overseas and visited numerous art galleries and museums. As a result, I have a long association with and interest in the visual arts both in Canberra and abroad. This unique upbringing, combined with my professional background and a dissatisfaction with contemporary collection interfaces, encouraged me to question how new Web technologies could be used to create more engaging methods of online collection access. Therefore the adoption of a practice-based research approach directly aligns with my own interests and values.

Practice-based research

Christopher Frayling has argued that creative production can be understood as a research method (Frayling, 1994) this view has been supported more recently by Estelle Barrett and Barbara Bolt (2007) who, in their book *Practice as Research, Approaches to Creative Arts Enquiry*, write that practice-based research has the potential to demonstrate how knowledge is revealed and how we come to acquire it. Drawing upon this concept by Barrett and Bolt, Mäkelä and Nimkulrat argue that:

> the exploration of knowledge partly through making artefacts has brought a new dimension to design research as the practitioner-researcher not only creates an artefact but also documents, contextualises and interprets the artefacts as well as the process of making them. (Mäkelä and Nimkulrat, 2011)

In addition to the processes of interpretation and contextualisation described by Mäkelä and Nimkulrat, it is important to also consider the role of reflection in the production of the creative works. I have been led by Schön's (1983) "reflective practice", where, through the process of reflection, I am able to build further knowledge of my practice. I will discuss these processes in further detail in Chapter five.

Joyce Yee (2010) draws attention to the potential of assemblage, or bricolage, as an approach which combines methods from various fields, in order to establish a suitable model of inquiry. Joe Kincheloe (2001) expands on this concept by writing that it enables researchers to "draw together divergent forms of research [and] gain the unique insight of multiple perspectives." In my project this was achieved by initially completing significant primary and secondary research—demonstrated through the practice and literature reviews—which situates my research project within a rich and divergent field of contemporary cultural practice.

In her review of practice-based research Candy (2009) argues that despite the concept of practice-based research having been around for some time "the existence of a discipline, defined by a body of theoretical literature, agreed methods and an archive of precedent theses is yet to be established." Other commentators agree that methodological innovation is required (Yee, 2010) as there are no commonly agreed research methodologies (Candy, 2009). Similarly, Malins and Gray (1995) hold the view that methodological approaches should adapt to changing paradigms of inquiry, which will in turn enable "a more complex, rich and holistic approach to research."

My research project has two interconnected components. I am concerned with both solving a problem: moving away from the search box as the primary method of collection access; and also encouraging exploration as an open-ended task. Rather than defining a single methodology to suit, I pick and mix different methodologies to "derive a suitable model of inquiry" (Yee, 2010).

Given that my project aims to encourage free-form exploration and serendipitous discovery, I did not consider the traditional task-based research methodology that dominates the fields of Computer Science or Human Computer Interaction (Rogers, 2004) to be an appropriate one. The limitations of this approach are clear in this example of an evaluation study of ArtVis, an application that combines visualisation techniques and tangible interaction to explore a large digital collection (Dumas et al., 2014). The authors write:

> The main goal of our evaluation was to assess whether users who work with the ArtVis application for the first time are able to find relevant information via the predefined visualisation requirements (Dumas et al., 2014)

In order to judge the effectiveness of the interface the researchers conducted a "task-orientated evaluation" which required each user to interact with the application and to answer 12 specific questions regarding the contents of the collection, for example,

> "What was the most active Italian city during the Renaissance?"
>
> "Explore the region where you were born. Have a look at the artworks that were produced in your neighbourhood. Write down the names of some artists. Did you know them already? During which period were they active?"
>
> (Dumas et al., 2014).

If a user is able to answer the questions then they have achieved their goal and the interface is judged as successful. However, simply by posing a question, the possibility of open-ended exploration has been diminished.

In light of this, and in the interest of addressing my research questions effectively, I adopted a number of different research methodologies that combined both theoretical and practical research components.

Summary

Having introduced the key research methodology that underpins my project, I will now describe the individual methods that I utilised. Following this, I illustrate how the braid provides a way to encapsulate the methodology and various methods and represent the relationships between them.

4.3 Methods

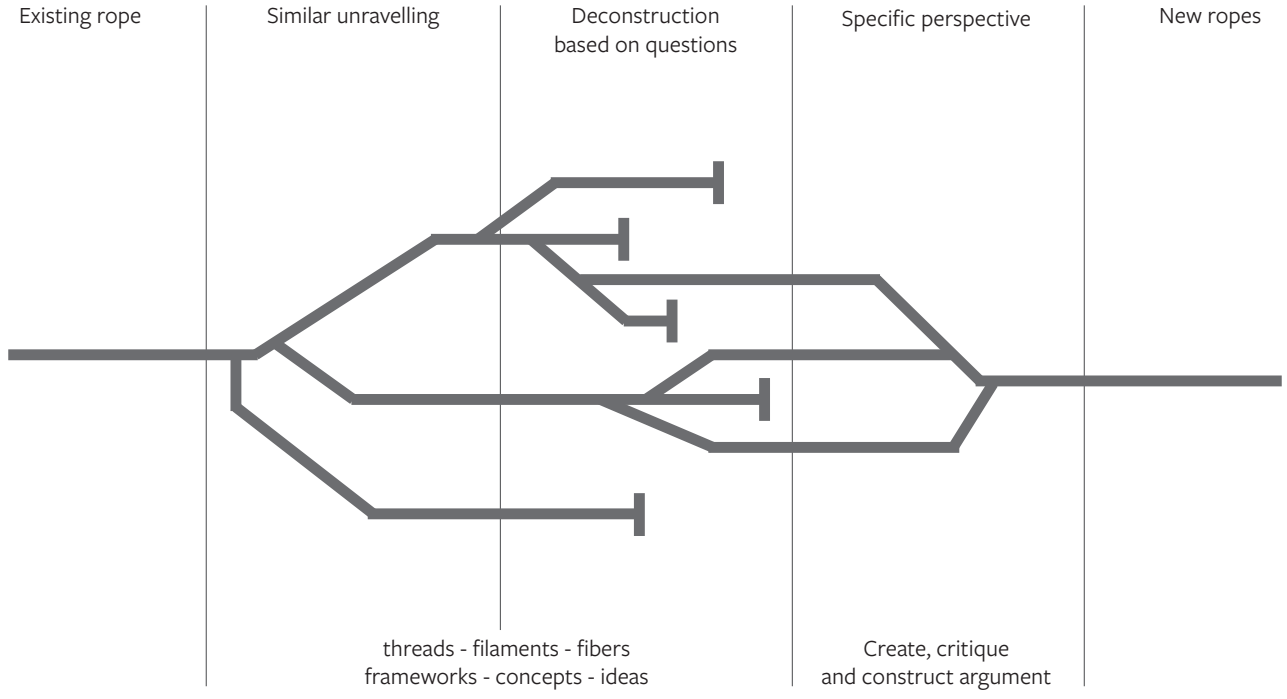
The production of the practice-based work was informed by a number of methods: the braid, sketching in code, primary and secondary research, prototyping tools, and finally, a mixed-method evaluation study, the results of which were used to inform production and provide reflections on the interfaces. These are listed in a logical order below, however, they do not sit separately and all inform each other.

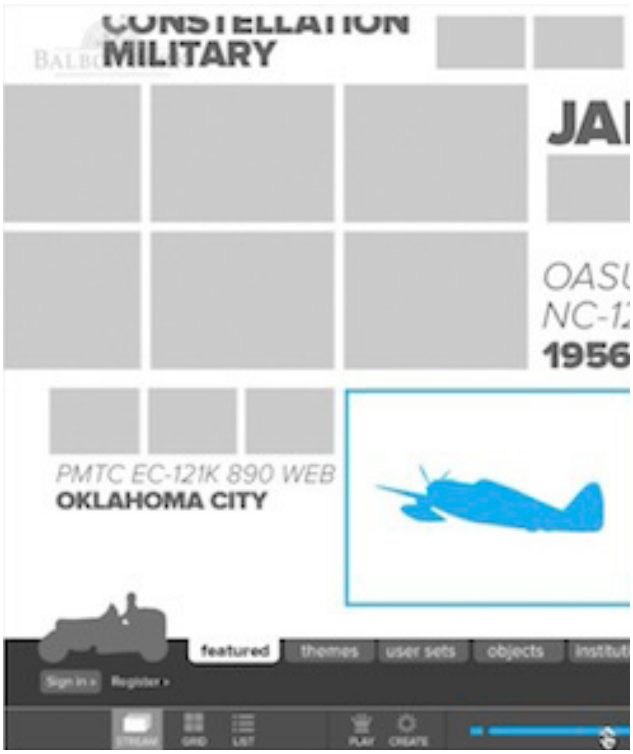
The braid metaphor

Graeme Sullivan's braid metaphor was particularly useful at the beginning of my project as it enabled me to conceptualise the entire research process (Sullivan, 2005). The braid (Figure 23) provides an overview of the various stages and demonstrates how the research develops as different ideas and concepts are explored. In the context of my research project the braid seems to acknowledge the messy and organic qualities of exploring the field, as well as the complexities of practice.

The braid or rope progresses through three stages. The first sees the rope starting as a whole before unravelling to reveal

Figure 23
Sullivan's braid





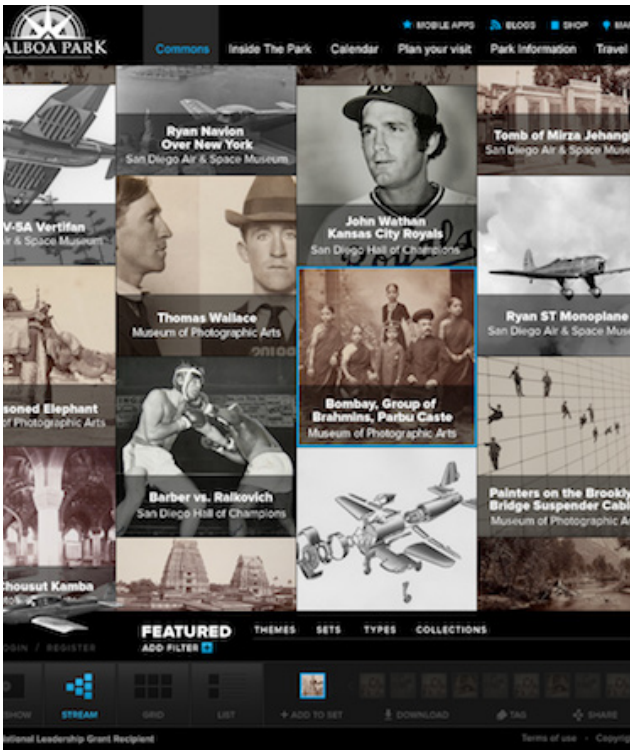


Figure 24 (left)
Wireframe for Balboa Park website (DePaolo et al., 2012)

Figure 25 (right)
Flat design for Balboa Park website (DePaolo et al., 2012)

the many different fibres or strands that represent concepts and ideas. In the second stage these fibres enter the deconstruction stage where questions are asked and answered—importantly some strands reach a dead end where they are no longer relevant to the line of inquiry and thus are discarded. Finally, the fibres enter the specific perspective stage, when new ideas are formed and arguments are constructed the different strands reunite and the rope becomes one again: a new rope which stands for the original contribution of the research as new knowledge.

In regards to my research project, the similar unravelling stage represented the process right at the beginning where I started to play with the data and test basic concepts. At the same time, I was immersed in literature from various fields as I developed a theoretical context for my work. In the deconstruction stage I was totally embedded within the project, and I identified two sub phases. The first of these consisted of the development of the interfaces: *Works and Networks*, *Decade Summary*, *All Artists* and *The Fader*. In the second phase, coinciding with my third year, I focused on the evaluation of the first three interfaces and analysed the evaluation results. In the specific perspective stage all of the ideas and

concepts were becoming one. I used the results from the evaluation to directly inform the development of the final two interfaces, *Subjects Explorer* and *Timeline*. It is from here that two new ropes are formed: one represents the knowledge produced from the creation of the interfaces, and the other the interfaces themselves.

The braid metaphor and the process it outlines are appropriate for this project, enabling its conceptualisation and encouraging questioning, exploration, analysis and discussion. I will reimagine the braid, and its role in my research project, at the end of this chapter.

Sketching in code

In practice, my work follows a pattern of development where the design and production of the interface occur simultaneously.

This is a departure from the traditional method of website development where a clear distinction exists between the various stages and the roles of those involved in that process (Robbins, 2012). Luke Reimer (2011) describes the process: a "documented outline of steps needed to be taken… in order to complete a Web design project." These include planning; design of static templates; development of interactive templates; launch and post launch. The roles of those involved are specific: a Web designer creates the static design before a front-end developer creates the HTML and CSS required to make them interactive and, in some cases, a Web developer is required to write JavaScript to process data or connect to a database (Coyier, 2013). In the field of online cultural collections this approach remains prevalent. It is demonstrated in the creation of the Balboa Park Commons, a website that brings together digitised collections content from 27 museums and cultural institutions in San Diego (DePaolo et al., 2012). DePaolo et al. (2012) describe how two flat designs were created after a five month wireframing process (Figure 24), the latter was "what the developers are currently using to build the site interface." (Figure 25).

Encouraged by the proliferation of Responsive Web Design (RWD)[1], a new development pattern has emerged which encourages designers to stop "making static image-based mockups in favour of designing with code" (Griffin, 2012). A code-first method is necessary for RWD, due to the dynamic behaviour of the page on different sized devices. Such an approach allows the designer to ensure the page works as intended at every stage of the production (Clarke, 2008), as they sketch directly in code, testing ideas and techniques in real time and resolving issues as they arise. This requires the designer (or developer) to master all aspects of the workflow, from layout and graphics to interaction design and data structure—a role also known as the 'full stack' developer.

Craft and Cairns (2009) argue that sketching should play a central role in design [and production] as it is a successful method for creativity and problem-solving, which has been central to the design-orientated disciplines in the past. David Verba (2008) agrees, writing that often the only way to really understand where problems are going to emerge from, or figure out a potential solution, is to sketch in code, so that the actual functionality can be prototyped. Koskinen et al. (2009) coined the term "protosketching" to define this approach. The process of writing code, testing (in browser), editing and testing again, borrows ideas from the iterative design methodology which is "specifically aimed at refinement based on lessons learned from previous iterations" (Nielsen, 1993). It provides constant technical validation as you know that you have something that works at every iteration.

1 First introduced by Ethan Marcotte in May 2010, http://alistapart.com/article/responsive-web-design.

In a relevant example from the domain of cultural collections, George Oates (2015) describes how she and Tom Armitage used this approach in the production of the *V&A Spelunker*[2], an experimental tool for exploring the Victoria and Albert Museum collection online. Armitage writes:

> The *V&A Spelunker* is an exploration of a dataset, but it's also very much a sketch—or a set of sketches—to see what playing with it feels like: not just an analytical understanding of the data, but also a playful exploration of what interacting with it might be like. (Armitage, 2015)

Armitage articulates the key idea that sketching allows the developer to understand the dataset itself and quickly prototype methods of interaction within it. It is impossible to predict or plan for complex features in an interface without first understanding the contents of the dataset and the complexities associated with it, including its structure, quality or consistency, scale and diversity. This process of data-exploration is an important one and can only occur in a code-based approach. It is not possible to sketch in just HTML and CSS, incorporation of the real data is an integral component of the approach. Armitage (2009) calls it "toiling in the data mines" and explains how even though the process can be overwhelming, it provides a better understanding of the data: what there is, how to represent it and what the core concepts are. Only by working with real data, in real time, is one able to design an effective interface based around it.

However, there are some weaknesses in adopting a practical approach that is essentially exploratory. Firstly, it is a complex process that requires the developer to have wide skill-set with a strong grasp of multiple programming languages—they need to be able to resolve problems that could be generated from many different sources. For example, the developer might need to write Python to output data through an API, then some HTML, JavaScript and CSS to investigate and explore the data before returning to Python to adjust the API call and modify the data returned from the database. Secondly, it can be

2 V&A Spelunker: http://va.goodformandspectacle.com

time consuming. The exploratory nature means that development is often unsystematic, with many obstacles that cause distractions or the pursuit of a solution which might ultimately be a dead end. Finally, there is a strong possibility of premature lock-in or succumbing to the sunk cost fallacy. As McRaney (2011) explains "your decisions are tainted by the emotional investments you accumulate, and the more you invest in something the harder it becomes to abandon it." In this case, because the process of development is difficult and time consuming it can be tempting to commit to the first idea, rather than discarding it and starting anew.

In my work, there is no substantial difference between the processes of sketching on paper to writing the code required to generate a Web interface. Both function like an initial draft that can be developed and refined through trial and error. I will provide reflections on this method in Chapter seven.

Primary and Secondary research

Primary research into the field was obtained by examining a wide range of online interfaces and analysing their functionality, aesthetics and technical development. Exemplars from this primary research component have been discussed in the practice review. Secondary research, as demonstrated in my literature review, shows how my work is situated within a varied field of academic research.

Prototyping Tools

In the production of the practice-based works, I utilised two different prototyping tools that were integral to meeting the intended outcomes of the research project. These are: Web design, through the context of the Web as a production and development platform, and the techniques it brings as a field of practice; and real data, which provides a form of built-in validation of the research project.

Web Design

From the outset, I envisioned how the Web would play an important role in the research project as a context for production. As well as being a development environment and a publication platform, it is also a scene of cultural practice. David Bell (2009) argues that the Web is simultaneously shaping and being shaped by culture, it is not "just a technological artefact, it is a cultural phenomenon, and these two 'realms' cannot be disentangled."

As mentioned previously, my professional background is deeply embedded within this field, both as a lecturer in Web design and development and as a commercial developer, therefore I was particularly interested in the creation of browser-based interfaces that use modern Web programming languages, specifically Hypertext Markup Language (HTML), Cascading Style Sheets (CSS) and JavaScript. This was preferable to relying on out-dated techniques, such as Adobe Flash[3] or a downloaded java applet[4]. I was also keen to experiment with the possibilities of client-side processing—where the scripts that help generate the page are executed in the browser—rather than server-side, where those scripts are executed on the server and then displayed.

Locating my research project within a Web-based context aligns it with a specific professional audience through the 'musetech' domain[5] and communities such as *Museums and the Web*[6] and *Museum Computer Network*[7], both of which run annual conferences presenting innovative research into digital practices in the cultural heritage sector.

Web design is a field of practice that is shaped by a diverse range of factors, from technological advancements to rapidly changing design trends. I use the term to refer specifically to the front-end aspects of a website or interface,

3 SFMOMA's ArtScope
4 CommonsExplorer and the Visible Archive
5 Musetech is an audience interested the application of technology in cultural institutions.
6 Museums and the Web: http://www.museumsandtheweb.com
7 MCN: http://mcn.edu

such as its aesthetic appearance: layout, typography and so on; as well as interaction techniques and data display.

The Web encapsulates a number of specific techniques and practices. I am interested in the development of modern standards-based interfaces which are written with succinct and semantic code. The W3C develops technical specifications and guidelines, known as standards, that are adopted by software developers to ensure the consistent display of content across different browsers (WaSP, 2006). As Web technology has evolved, so too have the various Web standards, which include HTML, CSS and JavaScript (Niederst Robbins, 2010; WaSP, 2006). New standards and specifications have been developed in response to numerous advancements in technology, for example, faster internet speeds; greater computer power; higher resolution displays and mobile devices (David, 2010). Over the course of this project, the HTML5 specification has transitioned from being partially adopted to fully adopted, and a number of CSS3 modules have reached recommendation level, including media queries in June 2012 (MDN, 2016). Modern standards-based Web development emerged as a response to the shortcomings of Microsoft's Internet Explorer, which failed to implement Web-standards in the browser, leading to much frustration within the Web development community (Finley, 2016). Modern standards-compliant Web browsers include Mozilla Firefox, Apple Safari and Google Chrome. By writing standards-based code, I can be confident that my interfaces will work on all modern Web browsers. In regards to semantic code, I refer to using the most appropriate mark-up to define the content on the page and ensuring separation between content, presentation and logic. I am committed to seeing how this will be applicable in the context of these data driven interfaces. As they are complicated Web based tools, I ensure the codebase is as clear and concise as possible, so that it is easy to maintain.

During the course of this research project many developments have changed the field. Particularly important is the continued rise of mobile and tablet devices, which increases ease of access and also brings a plethora of different screen sizes and

resolutions (StatCounter, 2016) to consider. This has resulted in a push towards Responsive Web Design—where designers use CSS media queries to ensure the Web page functions on different screen sizes (Marcotte, 2010); and alongside it, the emergence of User Experience Design (UX)—where a designer is tasked with ensuring the website is as user friendly as possible, so as to ensure the best possible experience (Unger and Chandler, 2012).

Hassenzahl and Tractinsky (2006) explain that while UX has recently gained momentum, the ideas it represents are not new. It is based upon the work of influential designer and academic Donald Norman (1988) who coined the term "user centred design" to emphasise the role of the user in creating a successful design. Norman (1988) argues that all things, whether an interface or a chair, need to be designed with the user in mind. Thomas Dickson (2008) agrees with Norman when he suggests that the experience a user has with something is based on its design, and these experiences can be pleasant, fascinating, surprising, provocative or repulsive. UX has come to be associated with a wide range of different meanings, for example, Norman and Jakob Nielsen (n.d.) define it as encompassing "all aspects of the end-user's interaction with the company, its services, and its products"; Luke Wroblewski (2009) sees it as a combination of information architecture, interaction design and visual design; and Law et al. (2009) describe UX as being "associated with a broad range of fuzzy and dynamic concepts, including emotional, affective, hedonic and aesthetic variables." It is clear that UX can be an abstract concept that covers a broad range of topics. Whilst I do consider many of these in my research project, I do not consider myself a UX designer, as my professional background and interests align with a much wider field of practice.

The importance of Web design as a research method is fundamental in the conception and production of the creative works within my research project. As identified in chapter 2, current museum collection access techniques tend to not leverage the power of the Web or embrace the possibilities of modern Web design. As a result, it is important that my work responds to these issues and considers the impact of Web design as a necessary component in an interface that aims to encourage exploration and discovery.

Real Data

The data that supports this project comes exclusively from the AP+P website, run by the Department of Australian Prints, Drawings and Illustrated Books at the National Gallery of Australia. This is a vast dataset that contains detailed structured data on more than 54,000 artworks; 19,949 artists; 3081 galleries; 8726 exhibitions and 9090 references. More than 26,000 of those artworks are imaged. Access to the data was provided through a partnership between the NGA and the University of Canberra (UC) that was initiated through a grant awarded in 2012.

I decided to focus solely on using this data for a number of reasons: firstly, and most significantly, the use of real data is a form of built-in validation of my research project. The interfaces produced are not merely proofs of concept, instead they are fully functioning, live and legitimate interfaces that are accessible online from the AP+P website. By using real data I am able to demonstrate that this is an authentic form of research. Secondly, the quality of the data is unparalleled in any other Australian collecting institution. It is an extremely rich online collection with strong connections between the various entities, for example, exhibitions are connected to galleries, which may be connected to multiple artists and so on. Finally, upon the conclusion of the Micro-Linkage grant, I was able to negotiate unrestricted access to the data.

Mixed-method evaluation

A traditional interface evaluation study requires a user to reach a specific goal within a controlled environment (Plaisant, 2004). Chris North (2006) explains that these types of evaluations are common because they enable the researcher to measure and compare results and are easily reproducible. However they are not without their weaknesses, as Plaisant argues:

> Finding and navigating tasks can be achieved without much difficulty by every [user], but effective exploration and discovery requires active intellectual engagement which is difficult to trigger and control. (Plaisant, 2004)

North (2006) agrees and highlights how this style of controlled experiment leaves little room for complex and relevant insight from users. He argues that the very nature of the experiment "forces users into a line of thought that they might not otherwise take" (North, 2006). It is clear that it would be difficult to measure how free-form exploration and serendipitous discovery could occur whilst using this type of research study. This problem is also identified by Pohl and her collaborators (2010) who remark that the exploratory nature of these new tools makes the evaluation process difficult and as a result alternative methods of evaluation are required that can capture the strengths and weaknesses of these new forms of interfaces.

From my prior professional experience, I knew that I would be able to obtain usage data automatically through the server logs, and if combined with feedback from users, I would have a wealth on which to reflect.

Creswell and Clark (2010) explain how a mixed-method evaluation study allows the collection of qualitative and quantitative data which will provide the most "informative, complete, balanced, and useful research results" (Johnson et al., 2007). Similar mixed-method research studies have been completed by Zhang and Marchionini (2005), Dörk et al. (2012) and Coburn (2016). Dörk and Coburn's research studies were also concerned with the evaluation of an open-ended Web-based visual exploration interface. Dörk et

al. (2012) discuss how a mixed-method approach provides more information both about human use and system performance, which, as visual exploration is a relatively new field, will lead to a better understanding of the field and its challenges. In these studies, the results are summative and used to validate the research project. However, in my project, I use the results of the mixed-method evaluation study to directly inform my creative practice.

It must be noted, however, that this style of evaluation study has one major shortcoming, which Jick (1979) draws attention to: "replicating a mixed-methods package… is a nearly impossible task." Despite this, he goes on to argue that it has vital strengths and encourages productive research (Jick, 1979).

Whilst many practice-based research projects do not have an evaluation process, it is not uncommon to include one. Candy (2009) describes how evaluation studies (within a practice-based research project) can be "beneficial, not only in providing substantive outcomes for the thesis, but also to the practitioner's understanding of the nature of their work." This view is supported by Rust et al. (2007) who explain that practice-based work can benefit from allowing evaluation "to follow on from the creative work, rather than relying principally on reflection in action." In other words, there are advantages in evaluating the work after its production, rather than completing it and then post-theorising it.

The importance of engagement with the user is a critical part of my work, therefore, in order to examine the relationship between the interface and the user I completed an intensive mixed-method evaluation study. This comprised think-aloud observations, a Web survey and associated data logging of usage. The results of the mixed-method evaluation study were used to inform production of *Subjects Explorer* and *Timeline*.

The mixed-method evaluation study is discussed in Chapter six.

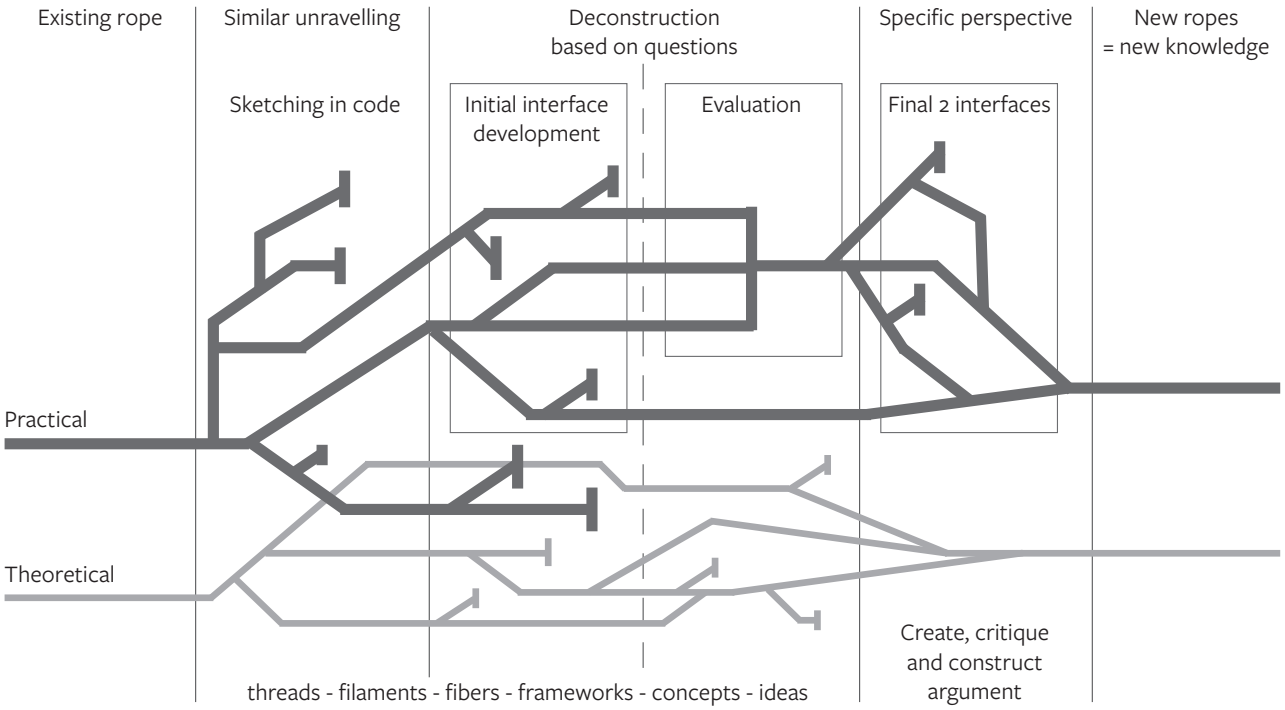


Figure 26
The braid re-imagined

4.4 Conclusion

At the beginning of this chapter, I described how Sullivan's braid metaphor provided me with a way to conceptualise the entire research process. I have adapted Sullivan's original graphic to show how it encapsulates the methodologies and methods I have utilised in my research (Figure 26).

This illustration provides a visual representation of my project. It shows how the braid recognises that the nature of practice is a complex one which requires a fluid approach in order to understand the field of knowledge.

In addition to the braid, I have also shown how the adoption of a further two methodologies, sketching in code and mixed-method evaluation, have allowed me to combine the necessary practical and theoretical components. I have described how a bricolage approach enables me to establish a suitable model of inquiry which brings together methods from various fields. Each of these methods: practice-based research, primary and secondary research, Web design, real data and mixed-method evaluation, inform the production of the creative work.

In the following chapter, I will introduce each of the interfaces produced in my research project. In Chapter six, I discuss the evaluation component and the results of a mixed-method evaluation study which sought to evaluate *Works and Networks*, *Decade Summary* and *All Artists*. In Chapter seven, I will outline my contributions to new knowledge and reflect on the practice-based works through a discussion that addresses my research questions in depth. Finally, in Chapter eight, I will summarise the project, its contributions to knowledge and discuss how the techniques I have developed are applicable to any digital collection.

5 The Works

5.1 Introduction

In this chapter I will introduce the six interfaces produced during my research project. I will discuss the development of each interface, its aims and design intent, before describing the outcome, the features of the interface and technical considerations.

The first three of these, *Works and Networks*, *Decade Summary* and *All Artists*, were developed in conjunction with my primary supervisor Dr Mitchell Whitelaw over an 18 month period between November 2011 and their launch in April 2013. Their creation was supported by a grant from the NGA and the Faculty of Arts and Design at UC. The production of the interfaces was a collaborative process; Whitelaw and I would often sit side by side to slowly sketch in the code. During this development period I wrote more of the HTML and CSS, while Whitelaw handled some of the more complex JavaScript and data processing scripts. It was a fruitful way of working as we were able to talk through ideas and problems and create the various functions required to develop the interfaces.

The next three interfaces were completed independently. *The Fader* was produced in late 2013 and was followed by *Subjects Explorer* and *Timeline* in 2015. The final two are the most substantial interfaces I developed, building on the earlier collaborative works, responding to the conclusions of the evaluation study and incorporating more complex technical processes.

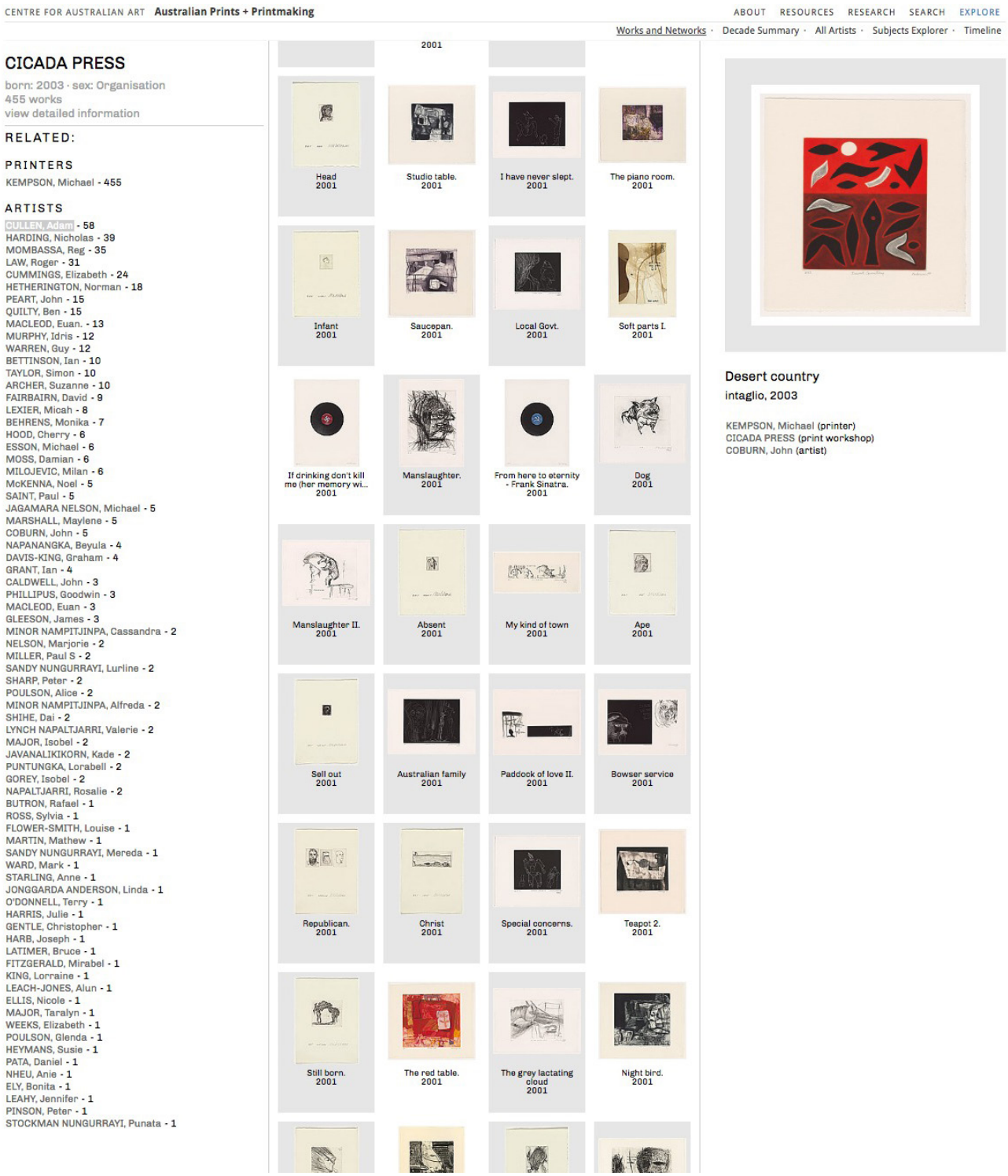


Figure 27
Works and Networks

5.2 Works and Networks

Available at:
http://printsandprintmaking.gov.au/explore/works-and-networks/

Development

Works and Networks (Figure 27) was the first interface produced as part of my doctoral research. Whitelaw and I knew that the AP+P dataset contained strong relationships between the various entities due to the collaborative nature of printmaking as a form of practice, for example, a single artwork could be connected to multiple creators. The aim of the interface was to explore these links. The name *Works and Networks* refers to the relationship between the artworks, the primary artist and other collaborators. Development of the interface began at the end of November 2011 and lasted for about five months.

Additionally, we used the process of production to come to an understanding of the complexities and type of data that existed in the AP+P dataset; how we could access it and, in turn, build a working interface with it—this is akin to Armitage's (2009) "toiling in the data mines."

At the outset, I framed my research project as focusing on data visualisation, however, as we started to develop *Works and Networks* it became apparent that relying on typical visualisation techniques would not allow us to achieve our objectives. This was made abundantly clear by an early attempt at a network graph (Figure 28) which attempted to visualise connections between artists or creators. Due to the phenomenal number of relations that exist in some cases, we regarded the visualisation as a failure because it was visually overwhelming. This encouraged us to begin experimenting with other ways to represent the data.

From a Web design perspective, our intention was to provide as much information within the single display as possible. This was influenced by my previous studies of other gallery and museum websites which showed that information tended to be presented

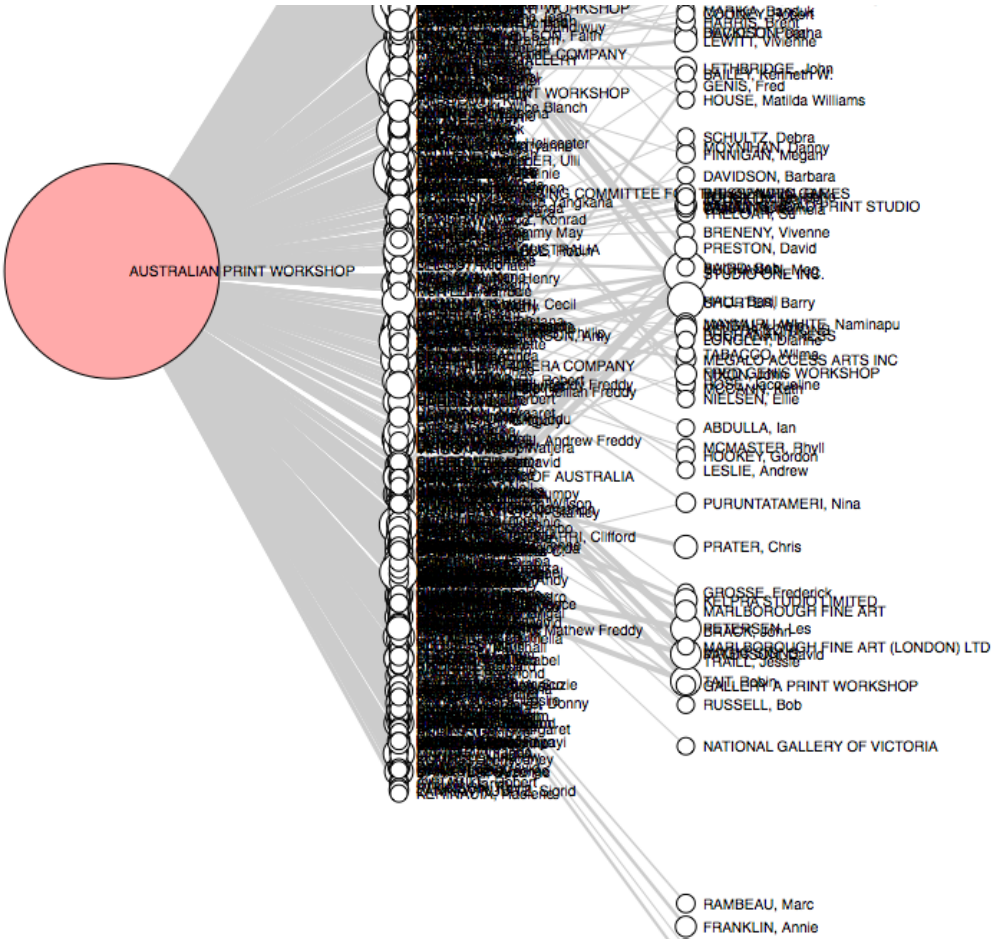


Figure 28
Early network graph

across multiple pages. For example, on the NGA Collection Search[1] interface, the following process occurs when looking for artworks by an individual artist: a search is formulated, the query returns a list of artworks, but more detailed information about each artwork is presented on a new page with a larger artwork image available on yet another page. Therefore, in order to view multiple artworks a user would need to return to the search results before proceeding to the next artwork and so on. This approach makes it impossible to see a visual overview of all the creator's artworks represented in the gallery collection. Nor does it allow the user to easily view an artist's artwork within a wider historical or comparative context. Figure 29 illustrates the NGA approach compared to the same process in *Works and Networks*.

1 See: http://nga.gov.au/CollectionSearch/Default.cfm

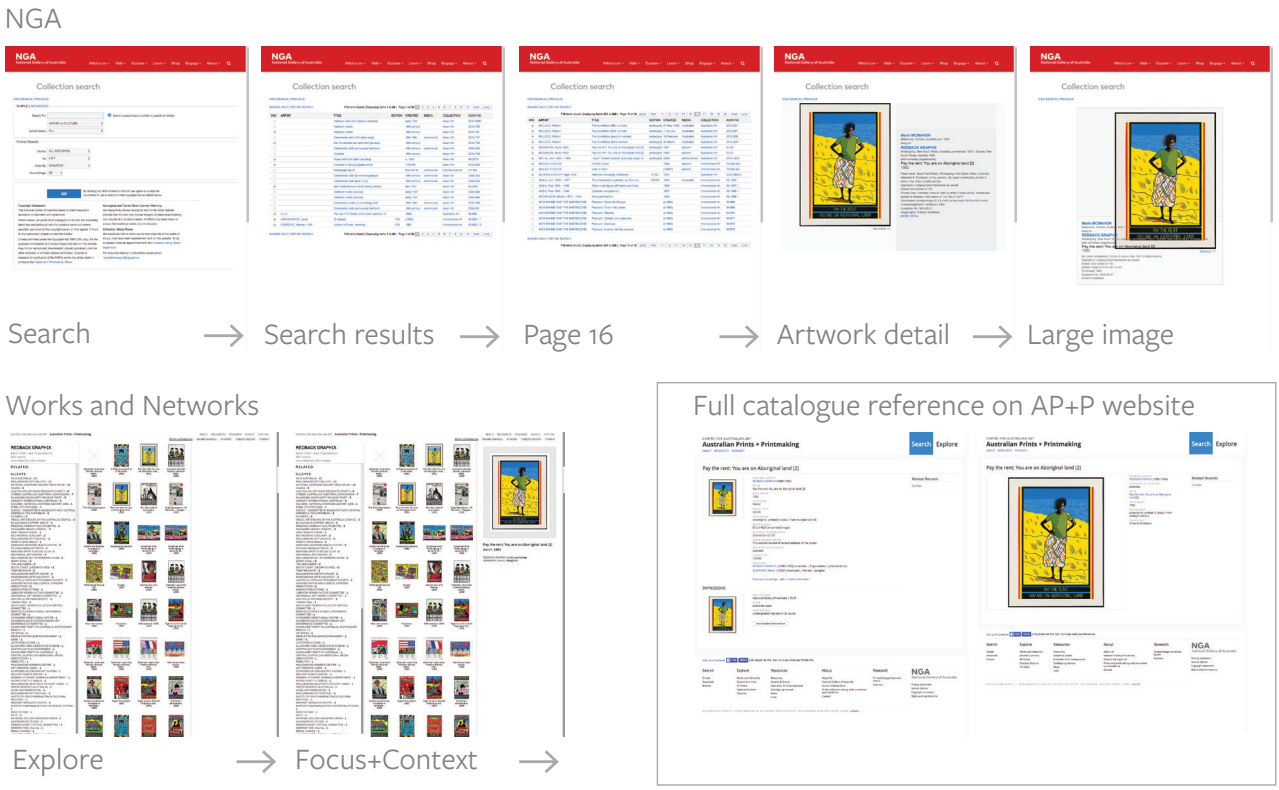


Figure 29
Flow diagram showing NGA collection search process compared to *Works and Networks* process

The NGA's Collection search interface has hardly been updated since it was first added to the NGA website in April 2008[2] and it does not consider how new technologies, such as wider screen sizes and faster internet connections, could support a more engaging and useful method of collection access. In fact, archived versions of the NGA website (Figure 30) show that the only substantial change made over the years was in late 2010 (Figure 31) when the site width was changed from 750 pixels to 980 pixels; even in 2016 (Figure 32) with the arrival of a 'new' NGA website, the Collection Search interface remains unaltered.

Outcome

Works and Networks uses a three-pane layout that provides multiple methods of accessing and viewing the collection data.

The left pane contains information about the selected artist or print workshop (their name, birth year, sex, artwork count and a permalink to their detailed record in the AP+P site) and a list of

2 The Wayback Machine first captured it in April 2008: https://web.archive.org/web/20080414083951/http://artsearch.nga.gov.au/

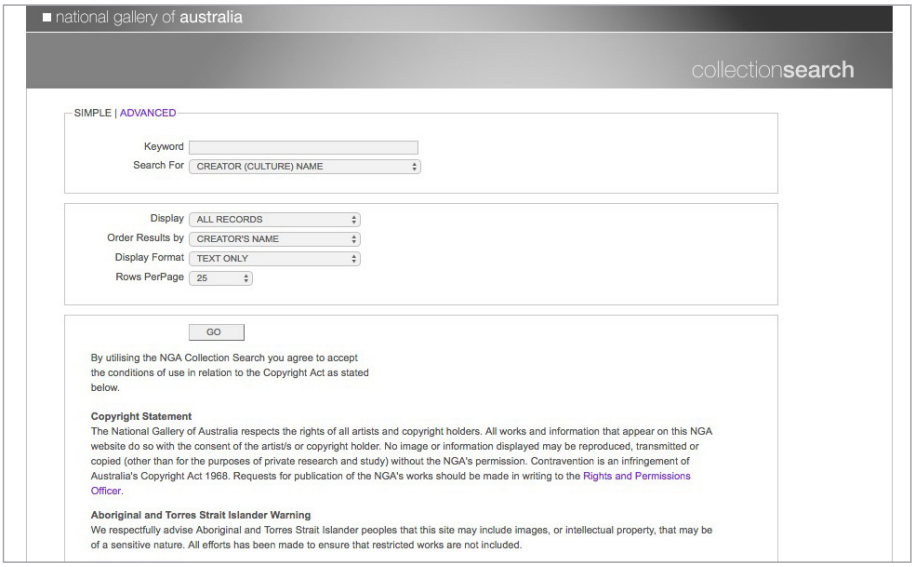


Figure 30
NGA Collections interface 2008

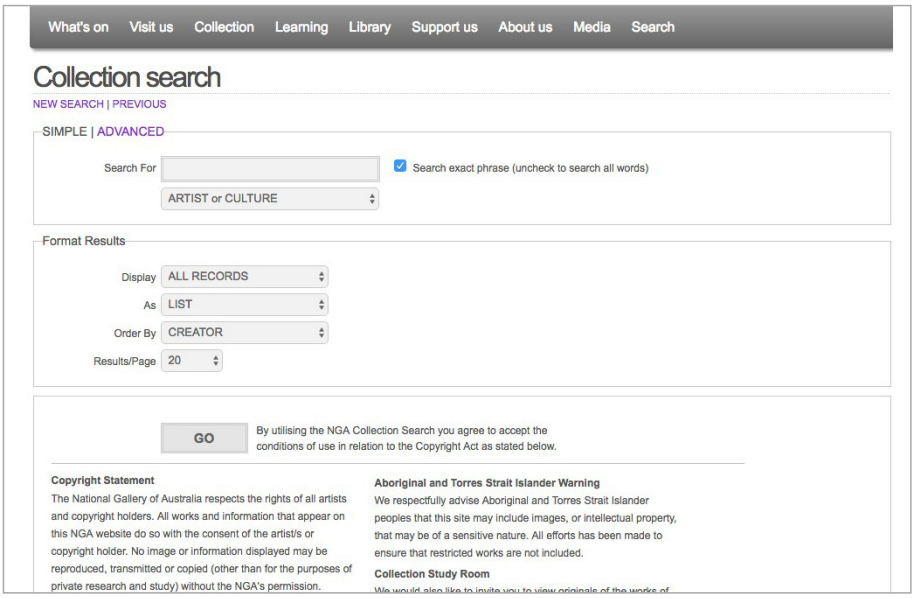


Figure 31
NGA Collections interface 2010

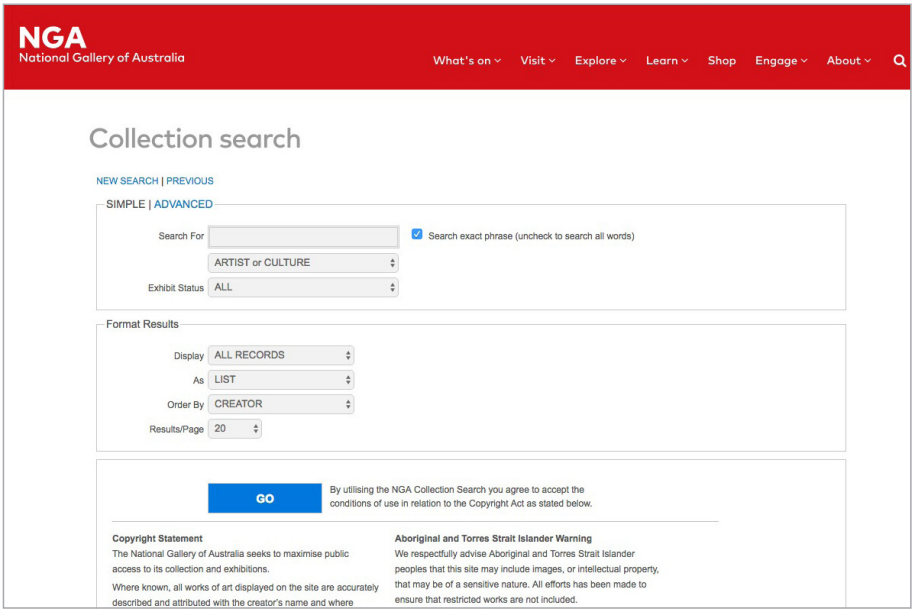


Figure 32
NGA Collections interface 2016

their related collaborators; the middle pane consists of a thumbnail grid of the artist's artworks, which includes the artwork title and date made. The right pane is empty until clicking a thumbnail, upon which a larger view of the selected artwork is displayed, alongside the title, date made, print type and full artwork creator information. This is a three level focus+context display, where it is possible to see the detail of the large artwork within the context of the other artworks and creator information. As the creation of a single print is not always a solo task multiple artists, printers, print workshops and so on, are often listed as additional creators.

The list of collaborators provides key functionality as it offers an opportunity to see who worked together and what artworks they produced. This list is generated by analysing the creators for every single artwork by the selected artist and as a result it represents the relationships that exist within each of those artworks. In order to visualise these connections more effectively, hovering over a name in the collaborators' list or hovering over an artwork thumbnail will highlight the linked artwork and related artists (Figure 33). This action reveals all kinds of relationships that might otherwise remain hidden. For example, the artist Basil Hall is listed as a printer on the artwork *Gara-Garag* by Queenie McKenzie[3], however this artwork also has an additional two printers and a print workshop attached; by hovering over these names in the left hand pane the user can start to form an understanding of Hall's relationships with different collaborators, including 57 artists, 3 print workshops and 18 printers. Clicking on a collaborator's name, either in the left pane or right pane, will refocus the *Works and Networks* interface on that artist, allowing the user to then see their artworks and collaborators. This interface encourages users to literally wander around the collection, from artwork to artwork, or from creator to creator, gaining an understanding of the complexity of an artwork's making and the relationships which underpin it.

3 Gara-Garag, 1997 by Queenie McKenzie: http://printsandprintmaking.gov.au/works/18120/

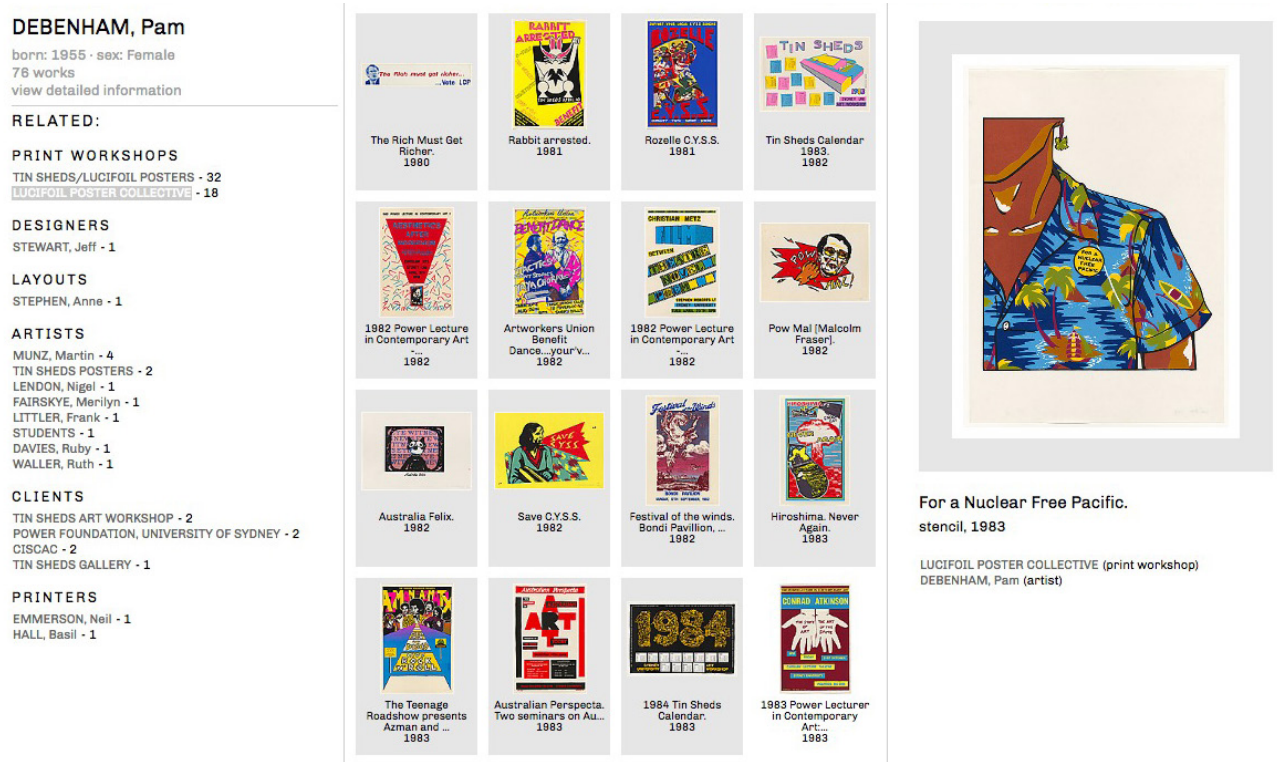


Figure 33
Three pane display in *Works and Networks*, showing highlighted relationships between creator and artworks.

The artworks themselves are displayed in a four-column grid, sorted by the date they were produced. Rather than excluding or hiding non-imaged artworks as in the Walker Art Center[4] and Rijksmuseum[5] websites, we chose to include all of the artworks so as not to distort the data set. Artworks with an unknown date are displayed first in the grid and have the abbreviation N.D (No Date). The size of the thumbnail was selected to ensure the full artwork remained uncropped and legible, thus maintaining the integrity of the artwork in its online representation.

It is important to ensure that a user is always able to access the "networked object" (Cameron, 2008) which is the detailed catalogue entry of an artwork or creator within the main AP+P website. Cameron (2008) argues that the "networked object" can take on a life of its own, where it is detached from the fixed place in time and space of the traditional collection site. *Works and Networks* contains two types of permalinks, a link on the large artwork image returns to the full catalogue entry on the AP+P

4 See Walker Art Centre default collection view: http://www.walkerart.org/collections/browse

5 Rijksmuseum: https://www.rijksmuseum.nl/en/search

site and by adding the selected artist id to the URL[6] we can ensure that each artist displayed within *Works and Networks* also has an associated permalink which can be bookmarked, shared or cited. The importance of the permalink[7] in this interface informed the generous interface principle to "share high quality primary content."

Accessing the interface through the standard URL[8] randomly loads a creator from a predetermined list containing the eight most popular print workshops—this ensures that the initial display contains enough data to display a rich overview of works by multiple creators. As there are many artists in the collection that have fewer than five works, this technique provided a useful starting point from which to explore. Access to the interface is also available via a link in the artist record page on the main AP+P site, or through the other *Explore* interfaces. In these cases, the interface will load with the focus on the selected creator.

From a technical perspective, *Works and Networks* is a dynamic data-driven interface. Rather than a static website where every page is pre-produced or a system where pages are generated on the server before being shown to the user, in our interfaces each view is created dynamically on the client-side. The data handling and injection of content into the HTML5 page is handled by jQuery, the base page is written in HTML with all styling being controlled by CSS. The data itself is retrieved from the AP+P database using custom API calls that were created for us by digital historian Tim Sherratt.[9]

6 For example: http://printsandprintmaking.gov.au/explore/works-and-networks/?artistid=8985

7 A permanent link that doesn't change, for example: http://printsandprintmaking.gov.au/works/46226/

8 Standard URL is: http://printsandprintmaking.gov.au/explore/works-and-networks/

9 At the time of development, Sherratt was also responsible for rebuilding a new AP+P database.

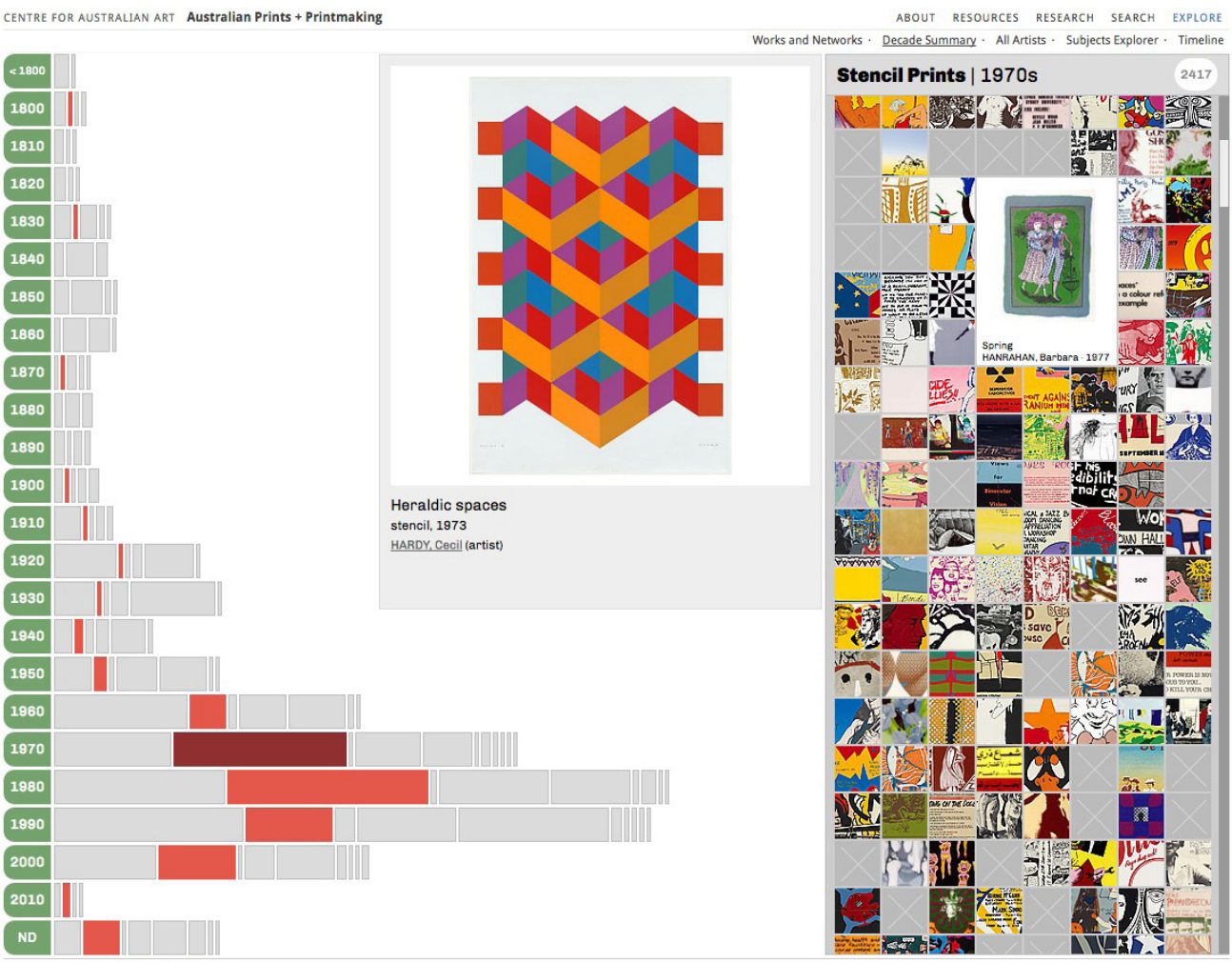


Figure 34
Decade Summary

5.3 Decade Summary

Available at:
http://printsandprintmaking.gov.au/explore/decade-summary/

Development

Decade Summary provides access to all artworks in the AP+P collection from one compact display (Figure 34). The interface encourages the user to play with the data by offering glimpses of colour and form in the cropped thumbnails and using the bar chart as an interactive navigational device.

Following the development of *Works and Networks* and the clear departure from utilising data visualisation techniques in that interface, we were keen to embrace these methods in *Decade Summary*. We suspected that we could combine classical data visualisation practices with pure HTML, CSS and JavaScript to create unique methods of viewing the data, instead of relying on a visualisation library, like D3[10], to produce a graph or network diagram.

The concept of the bar chart was influenced by work visualising the AP+P collection I produced whilst completing Honours at UC in 2010. In this project, I created a set of printed posters where a vertical bar chart represented the artworks produced in each year (Figure 34). Each column in the chart was made up by images of the artworks themselves. The project had shown that this data visualisation technique had plenty of potential but serious limitations emerged in regards to scale and legibility with a collection of this size.

We were also keen to develop the concept of an overview, as identified by Shneiderman in his Visual Information Seeking mantra. Whitelaw had previously experimented with the creation of overviews in his *Visible Archive* project for the National Archives of Australia and this style of access had proven to be very effective. Additionally, the strength of an overview as an access tool had emerged through my examination of a number

10 D3 is a JavaScript library for creating data visualisations: https://d3js.org

Figure 35
Detail of timeline of artworks sorted by decade produced as part of my Honours project in 2010

of interfaces (such as *ArtScope* and *Preservation of Favoured Traces*) and texts in the practice and literature review.

Development of the *Decade Summary* began in June 2012 and took 9 months.

Outcome

The interface embraces two main components: in the left pane is an interactive chronological bar graph showing the distribution of artworks by decade; in the right pane is a grid of cropped thumbnails of the artworks from the selected decade and printmaking process.

The horizontal bar chart is a classic data visualisation technique; it is useful in this context as it allows the display of a large quantity of information in a small and easily understandable manner (Figure 36). Each horizontal bar aggregates artworks from a single decade. Decades were selected because they gave a useful number of segments given the time span of the collection, and are a legible time period we are familiar with from references to the 1920s or 1970s in a cultural context. Along the left edge of each bar is a year label and the bar itself is segmented by the key printmaking processes represented in the collection: intaglio, stencil, monoprint,

planographic, relief, electrostatic, photographic, pigment transfer, computer generated and paperwork. This segmentation technique allows the user to see how the distribution occurs and changes over each decade. For instance, it reveals the popularity of stencil printing (also known as screenprints) in the 1970s and 1980s and the emergence of the computer-generated print in the 1990s. There are other valid ways to segment each group of artworks, but this technique is both practical (as there are a relatively small number of different values) and interesting (it shows distinct changes across the collection). The last bar contains those artworks with unknown dates; as with artworks with missing images, it is important to ensure they remain in the interface so as not to skew the representation.

The currently selected segment has a dark red background and a lighter shade or red indicates a connection with the corresponding segments in the other decades. Moving the cursor over a different one highlights the relations across the bars and reveals a label with the name of the printmaking process and the number of artworks within that segment. Clicking a segment loads the corresponding artworks into the right pane of the interface. This action, and the interactive bar chart itself, are both focus+context displays.

Figure 36
Detail of *Decade Summary* bar chart

Cropped thumbnails in the right pane provide an immediate visual representation of the artworks from the selected decade (figures 38 and 39). This technique built upon Hinton and Whitelaw's approach in *commonsExplorer* (2010) and involved much discussion and experimentation: we debated the best way to show the artworks in bulk, wondering if it was appropriate to crop the images; and testing how small we could crop them whilst retaining their legibility. Ultimately we decided that the cropping allowed a condensed display, with the thumbnails essentially becoming an abstract visualisation in their own right. Each offers a hint of colour, form and texture that provides a good representation of the contents of the artworks. This in turn encourages the user to interact with the interface and see the full artwork. Hovering over the cropped thumbnail reveals an uncropped image with the artwork title, primary artist name and date made. Clicking on a thumbnail will load a larger artwork image within the interface—another focus+context action.

We were fortunate that the quantity of prints in the collection from the early nineteenth century was quite low, resulting in a large empty space in the centre of the interface. This allowed us to add a container in which a user could view a larger artwork within the context of other artworks shown in the interface. By having a strong understanding of the data, we were able to utilise the unique properties of the collection to facilitate the user experience. The large artwork view displays the title, date made, print type, any listed creators and their roles. We reuse the method from *Works and Networks*, where clicking on the artwork image returns to the catalogue entry on the AP+P site and a link on the creator's name opens a focused *Works and Networks* browser, see figure 37.

As with *Works and Networks*, loading the *Decade Summary* from the direct URL[11] will load a default starting point. In this case, the selection is already on Planographic Prints from the 1960s—chosen to provide the user with a hint of how the interface works—without explicitly providing instructions.

11 Standard URL is: http://printsandprintmaking.gov.au/explore/decade-summary/

The *Decade Summary* follows the same technical approach as *Works and Networks*, with the page being generated dynamically, using jQuery, HTML5 and CSS. The data is retrieved from a number of custom API calls developed specifically to generate this interface. In order to reduce load times, on both the server and the users computer, the first call provides a summary of the collection data as a whole, which is used to generate the bar chart:


```
{
      "decade": 1790,
      "total_artists": {
            "gender": {
                  "Female": 1,
                  "Male": 13,
                  "Organisation": 0
            },
       "total": 15
      },
      "total_works": {
      "media_cats": [{
            "media_cat": "Print",
            "total": 32
            },{
            "media_cat": "Illustrated book",
            "total": 19
            }
      ],
      "print_types": [{
            "id": 1,
            "print_type": "intaglio",
            "total": 49
      },{
            "id": 5,
            "print_type": "relief",
            "total": 1
      }],
      "total": 51
      },"url": "http://printsandprintmaking.gov.au/
dvapi/summary/decades/1790/"
}
```


A second call returns artwork information for the selected decade and print type, which creates the thumbnail grid. Instead of making a third API call to return individual detailed artwork information (to display the large artwork view) we load the image separately and reuse the data returned in the second API call.

By providing a permanent link back to the full catalogue entry for an artwork on the main AP+P website, we recognise that users need to be able to bookmark, share and cite the different views within

Figure 37
Detail of tool-tip showing hint to open artist in different interface

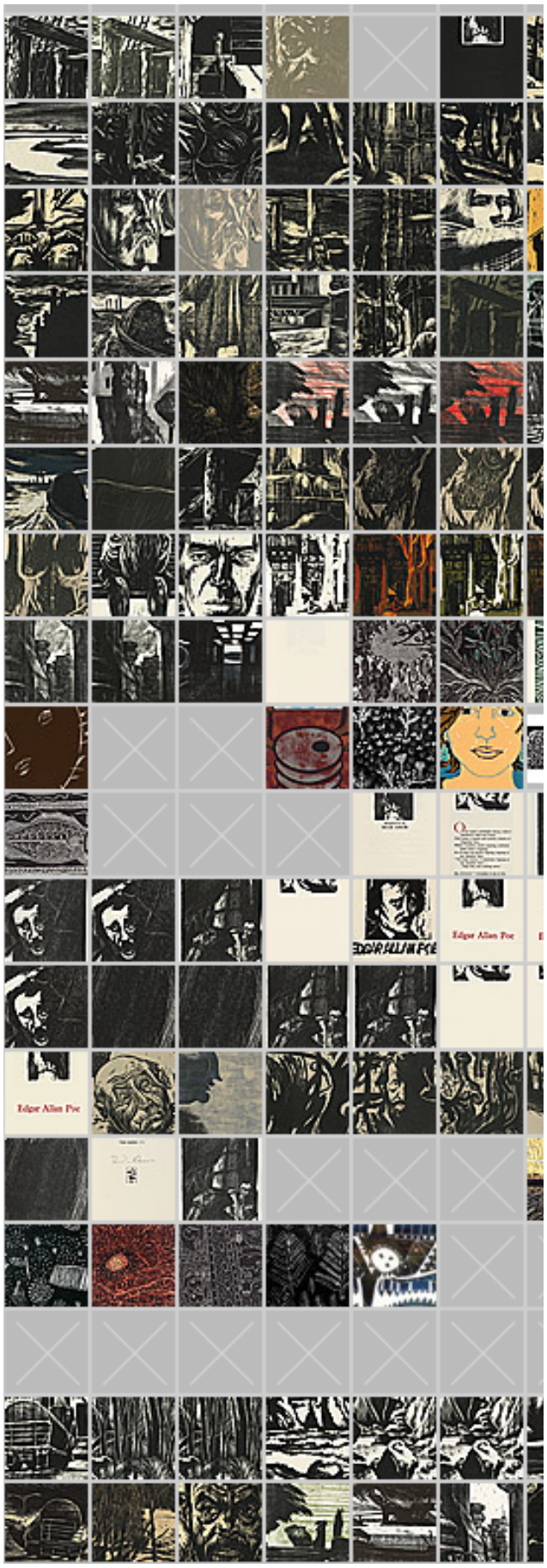
an interface, for example, if they are viewing Intaglio Prints from the 1820s. However, the problem with this style of data-driven single page application (SPA) is that all interaction occurs within the same page, so the URL never updates. To overcome this issue, we use the hashchange event and append hash fragments[12] to the URL as the user clicks on the various segments in the interface. For example, the URL when initially loading the interface is:
http://printsandprintmaking.gov.au/explore/decade-summary/,
if the user were to navigate to Intaglio Prints from the 1840s, we add a hash, the year (1840), and the id of the printmaking type (1):
http://printsandprintmaking.gov.au/explore/decade-summary/#/1840/1
This ensures that a user can return to the view at anytime or, as the hashchange event uses the browser history API, each state is added to the history which enables back button navigation.

Figure 38 (left)
Detail of cropped thumbnails of planographic prints from the 1850s

Figure 39 (right)
Detail of cropped thumbnails of relief prints from the 1990s.

12 JavaScript hashchange reference
https://developer.mozilla.org/en-US/docs/Web/Events/hashchange


FALMOUTH


Edgar Allan Poe
Edgar Allan Poe

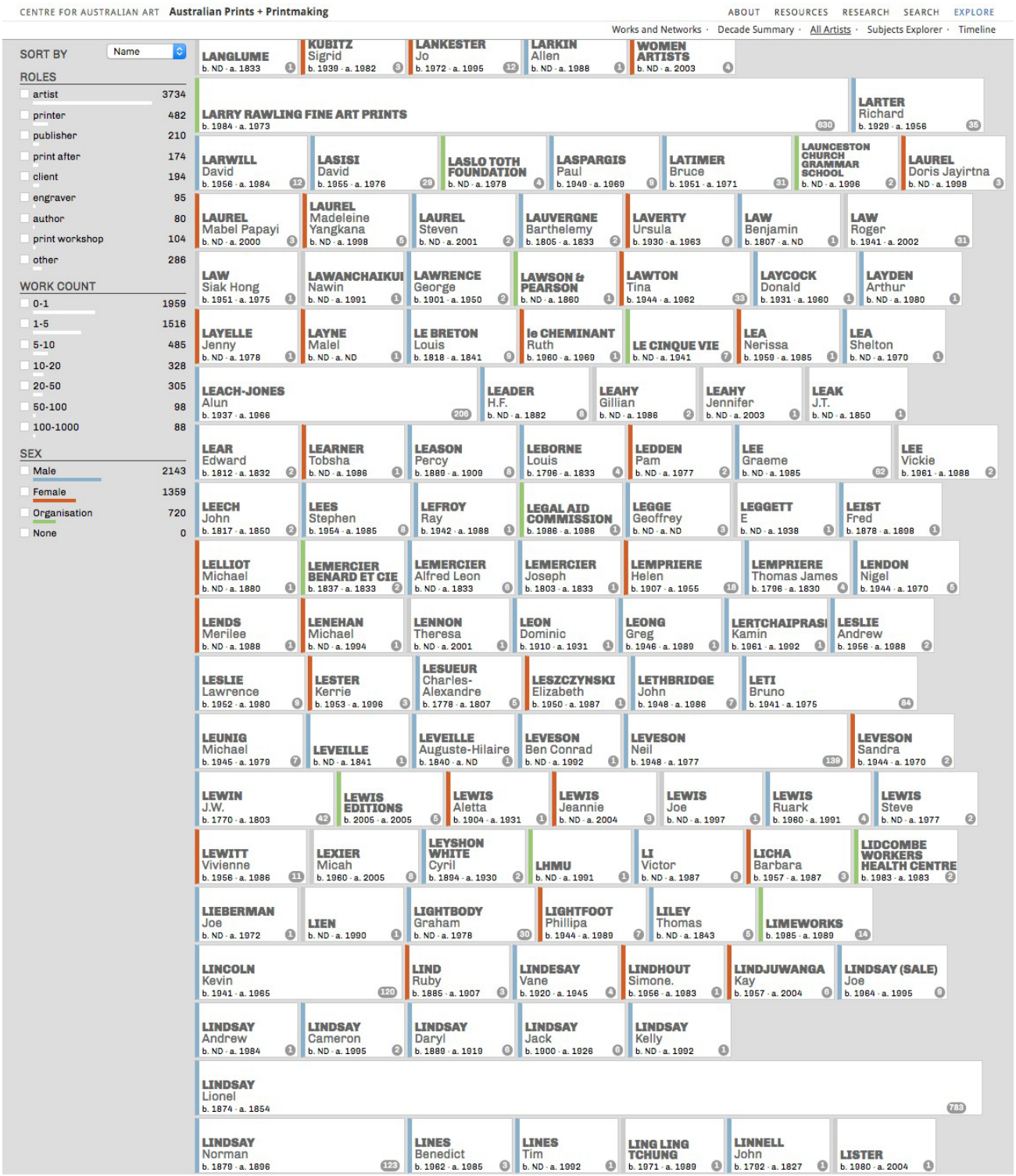


Figure 40
All Artists

5.4 All Artists

Available at:
http://printsandprintmaking.gov.au/explore/all-artists/

Development

The *All Artists* (Figure 40) interface provides a rich, artist-focused overview of the collection. Its aim is to encourage exploration of the individual artists' artworks in depth. It introduces new techniques that allow a user to compare multiple artists and their artworks, and to filter the data by different facets within a single contextual display. Unlike the previous two interfaces, there are no preselected segments or predetermined artists, instead all the data is shown at once and the user makes their own decisions about what to view.

The interface was the final one developed as part of the original grant from the NGA and UC. It was produced over a four-month period between March and June 2012.

All Artists started with a simple question: *was it possible to create an interactive containing every single artist within the AP+P collection*? From this starting point, it evolved into a truly experimental prototype. We aimed to test the limits of the browser by constantly improving our sketches and asking further questions: could the browser actually draw more than 5000 different sized boxes on the screen at once? What if the user could filter this data in real-time? What if these boxes were interactive? Would the browser crash? And so on. The outcomes of this experimental approach are discussed below.

Outcome

In this interface, 5359 boxes are displayed on the screen at once, each of which represents an artist that has at least one artwork in the AP+P collection.[13] Each box contains the artist's name, year of birth,

13 Unlike a normal gallery collection site, the AP+P website contains information about many artists who are not represented in the collection—we wanted to focus on those that have works as that indicates their connection to the actual NGA collection.

details of when they were active and the number of their artworks in the collection. The individual boxes have a minimum width which is extended to reflect their total artwork count. Printmaking can be a collaborative process, which means that some artists have multiple roles (for example, a printer or binder) that might have a higher artwork count. A coloured border on the left indicates their gender: blue is male; orange female; green for an organisation and grey is unknown. These simple data visualisation techniques enable the user to quickly scan through the interface and see the distribution of different genders and gain an immediate understanding of the number of artworks by each creator in the collection. The artist boxes are initially sorted by last name and a drop down selector on the left allows re-sorting by birthdate or active date. The initial display of all the artists influenced the Generous Interface principle to 'provide rich overviews', additionally, it is the first step in Shneiderman's Visual Information Seeking mantra.

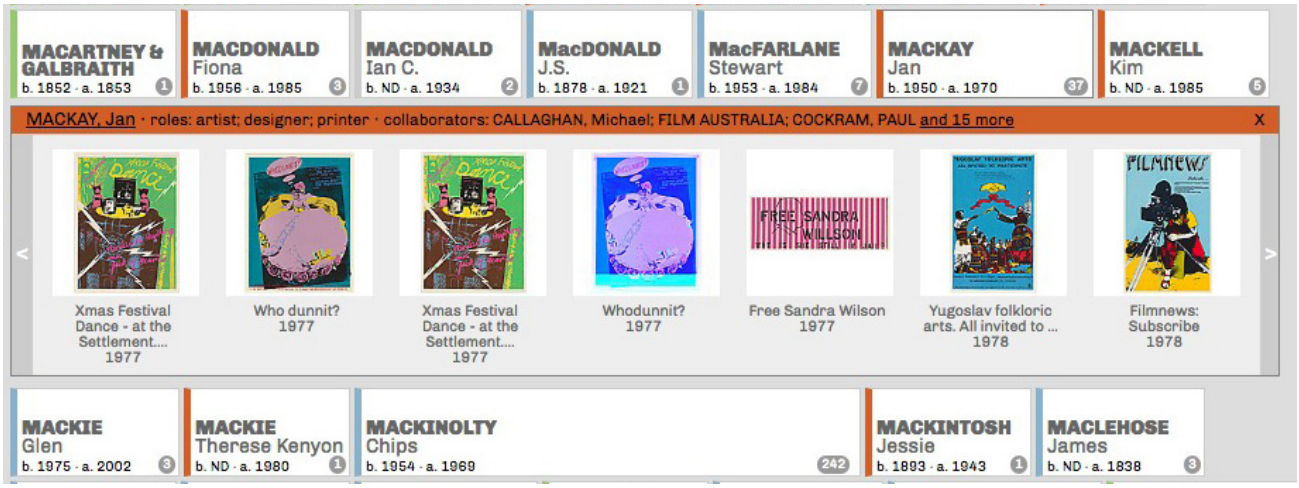


Figure 41
Detail of an opened artist container in *All Artists*

Clicking an artist's name will insert a new full width container in the row immediately below the artist box (Figure 41). This container provides access, within the interface, to images and details of their artworks, their listed roles and any collaborators who were involved. Importantly, a user is able to open multiple artist containers at once, which enables comparison between artists and their artworks, without having to move to a new window. For example, a user can concurrently compare the lithographs of Deutsch & Ferguson from 1859 to artworks produced by

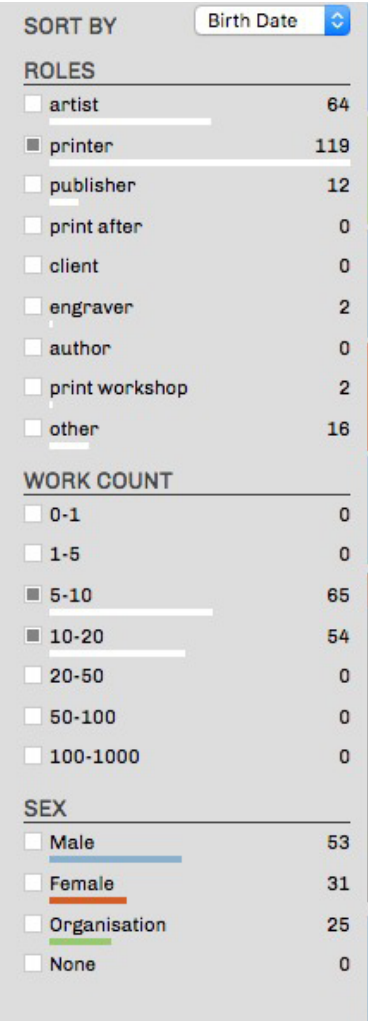


Figure 42
Detail of selected facets in *All Artists*

LITHOS Press in 1985. This is a great example of a focus+context display and a method for simultaneous comparison, both of which I develop further in *Subjects Explorer* and *Timeline*.

In order to deal with the size and complexities of the dataset, a number of filters are provided on the left of the interface (Figure 42). These were inspired by the faceted filtering that occurs in online shopping[14] and they allow the user to quickly refine the data displayed on-screen in real-time. This is a common convention in search result listings and is partly influenced by Marchionini's concept of exploratory search. For example, by selecting the checkbox for 'printer' and 'female' the boxes on screen are instantly updated to only show the 112 female printers. At this point the filter count has also updated to show that of these, 30 printers have 1-5 artworks and only 4 have 100 or more artworks. Underneath the title of each filter is a horizontal bar that allows an immediate visual understanding of the distribution, I call this technique quantitative aggregation, and I will discuss it further in Chapter seven. At all times it is possible to open an artist container and view their artworks. This means a user can quickly narrow down the information displayed and show only what is relevant to their interests or needs.

As with the previous two interfaces, *All Artists* was written with jQuery, HTML5 and CSS. A custom API call returns the artist data in blocks of 500 results at a time, these are displayed as they are processed with a loading icon on the top left. Attempting to retrieve all the results at once would exceed the limits of the server and cause it to timeout or crash. This delay in retrieving the data means the user has to wait until all of the data has loaded to ensure they are able to see the full overview and use the filters. A second API call is made when a user clicks an artist's name, which returns individual artist information and their artwork data.

14 For example on ASOS: http://www.asos.com/men/bags/cat/pgecategory.aspx?cid=9265

397 works by HANRAHAN, Barbara Play Pause Change focus to decade / type

← →

not titled ↪

1962 · relief print

HANRAHAN, Barbara · artist · 406 works

Figure 43
The Fader

5.5 The Fader

Available at:
http://printsandprintmaking.gov.au/explore/the-fader/

Development

The Fader was conceived as an additional entry point to the collection, one that would be enjoyable, immersive and slow paced (Figure 43). It was inspired by the photo slideshows that used to dominate personal computer screensavers[15]. It is in fact a performance of the collection that takes a very different approach to the previous *Explore* interfaces. It was the fourth interface produced during my research project and the first completed upon the conclusion of the grant and without Whitelaw's contribution.

Instead of providing an overview of the collection *The Fader* presents a focused view of individual artworks from either a decade and print type or a single artist. It starts with a singular artwork being presented on the screen and encourages the user to simply observe the collection as one artwork slowly fades to another, with no hint of what will follow. It is a surprising display, which even in testing revealed many previously unseen artworks from the collection. *The Fader* also gave me a platform to experiment with new Web tools.

This interface was designed to function on an iPad as well as a standard computer screen. At the time of production in 2013 there was a substantial push within the Web development community towards Responsive Web Design.[16] As a result, I was keen to test how these kinds of data-rich interfaces could work on mobile devices, whilst also encouraging more informal exploration options.

The earlier interfaces had many different elements: lots of text, graphics and layout components, and for *The Fader* to work, it needed to be different. Aesthetically, I wanted to create a minimal design that could allow for an immersive

15 For more on 'Screen Saver Culture', see Watz, 2012.
16 See Google Trends report on searches for 'Responsive Web Design': https://goo.gl/LiamxQ

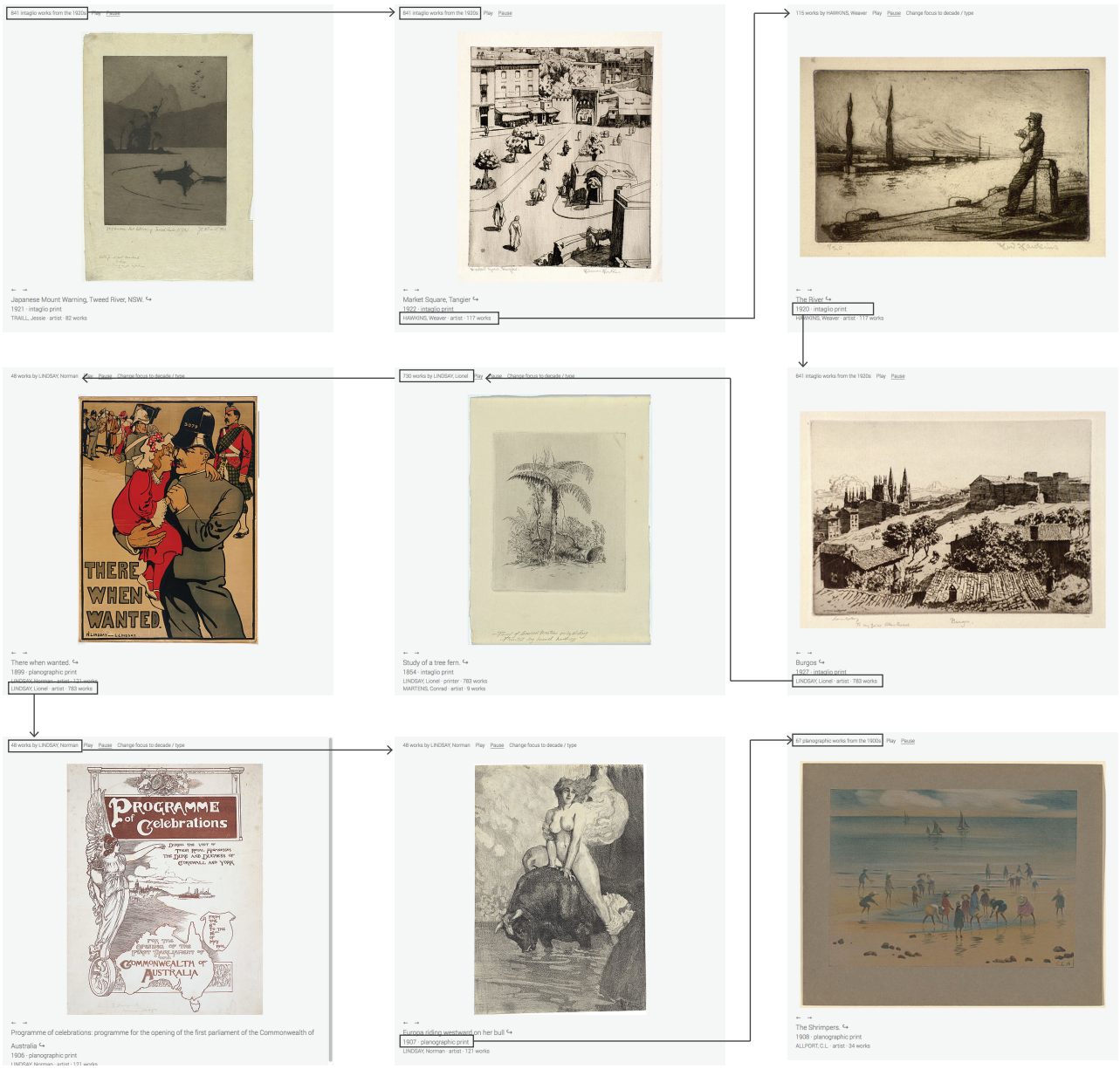


Figure 44
Spidering process in *The Fader:*

experience, but one which incorporated enough basic detail for the user to understand what they were viewing.

Outcome

In *The Fader* the display is split into three rows. The first contains a label describing the current focus of the interface, followed by play and pause controls; in the second row a large artwork image is displayed (in this interface the data is filtered to remove any non-imaged artworks); and the final row contains additional navigation controls as well as basic artwork information (title, date, type and artists details).

Navigation within the interface allows the user to immediately move to the next artwork by using the previous and next arrows positioned above the artwork title; clicking on the artist's name refocuses *The Fader* to only include artworks by that artist. If this occurs, an additional label appears on the top row: 'Change focus to decade / type', which in turn allows the user to pick a different decade and print type to focus on—slowly spidering across the collection (Figure 44). For example, when loading the interface, the initial focus is on relief prints from the 1960s, as *The Fader* cycles through the artworks different artists' names are shown—clicking on one of these names, say Weaver Hawkins, will refocus it to cycle through only his 115 artworks with images. Many of Hawkins' artworks are intaglio prints from the 1920s, and so, clicking the date and type will change the focus of *The Fader* to the 841 intaglio prints from the 1920s.

To support navigation on an iPad, I borrowed a convention from Facebook where the user swipes the image to navigate through the artworks.

Further testing has shown that *The Fader* works particularly well on large format display screens, such as a television or multi-screen display (Figure 45).

The Fader was originally built using the same jQuery, HTML5 and CSS approach as the earlier interfaces. However, after some experimentation with the open-source Web framework AngularJS

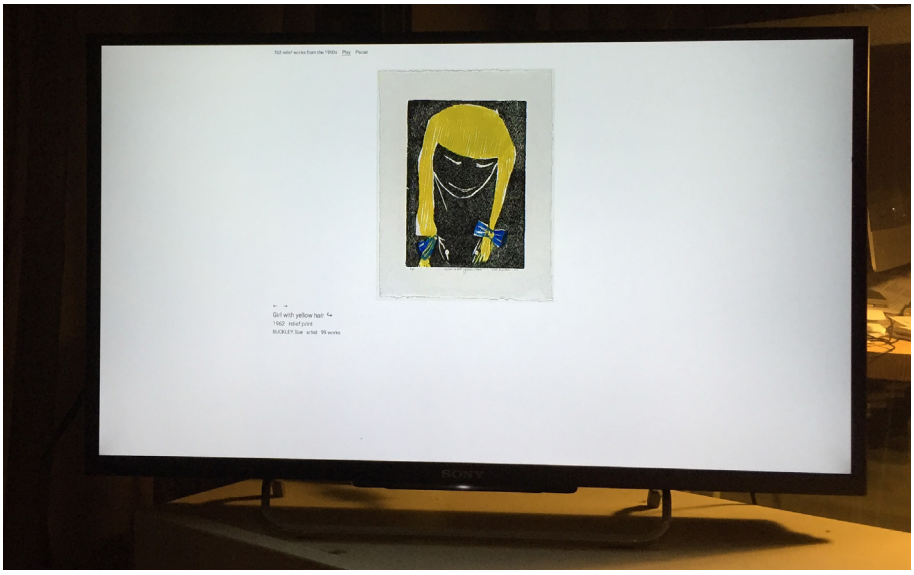


Figure 45
The Fader running on a television

Figure 46
Landscape artwork displayed in *The Fader*

(hereafter referred to as Angular) I decided to move away from jQuery and rebuild it using the Angular framework. There were two main reasons for this: firstly, Angular embraces the Model-View-Controller (MVC) framework which separates presentation, data and logic from each other, in contrast to jQuery which does not require such separation. Secondly, it has two-way data binding which automatically updates the view whenever the model changes and vice versa, making the process of binding the data to the HTML much easier. In other words, the HTML document contains references to the data source, for example, `{{ artist.name }}` might define an artist's name. When the page is loaded, the data model is updated and the artist's name is inserted into the HTML (the view), if a user were to click on a different artist, then the model would be updated again, and the new artist's name would be inserted into the page. All of this occurs without the page ever having to be reloaded, which provides a much smoother interaction experience for the user.

The rebuilt Fader allowed me to test whether such a framework could be used in the production of additional interfaces. In regards to the data, rather than producing a new custom API specifically developed to generate this interface, I show that is possible to reuse API calls from earlier works: one call comes from *Decade Summary* and another from *All Artists*. This shows that very different interfaces can be created using the same data source, and that an interface of this kind can be produced without requiring custom APIs.

Production of the interface took place in the last months of 2013, and whilst it is accessible on the AP+P server, it was not publicly released. This was primarily due to the evaluation study already being underway at the time of the completion. Later, after discussion with Whitelaw and Butler (at the NGA), I decided not to release it publicly due its departure from the 'overview' style of the other interfaces.

It is remarkably different to the other interfaces as it moves away from visualisation or overview techniques and reveals a wide range of more evocative interfaces or displays which could be investigated in future work. *The Fader* could be considered as a 'performance' of the collection, rather than a visualisation. The development of this interface was a rewarding process as it allowed me to experiment and test various techniques and new tools which were greatly beneficial in the development of the last two interfaces. The adoption of the Angular framework proved to be the most rewarding of these experiments as it perfectly suits the data-centric style of my work.

Australian Prints + Printmaking
Subjects Explorer

Selected subject:
Australia: Travel Poster has 7 related subjects

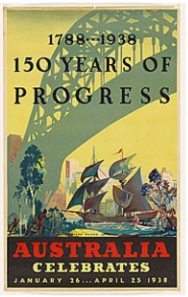
Selected filter:
Charles Troede ... ×

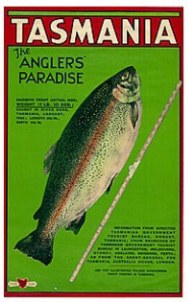
- Australia · 6927
  - Book Arts · 3071
  - Art Style · 1394
  - Printed Ephemera · 1178
  - Art Period · 976
  - Animal · 333
  - Stencil Art Collection · 286
  - Plants · 178
  - Multicultural · 166
  - Gold Rush To Industry · 136
  - Political Figures · 74
  - Travel Poster · 73
  - Bicentennial, 1988 · 68
  - Gold Rush To Industry 1851-1901 · 35
  - New South Wales · 14
  - Elections · 13
  - Sport · 10
  - Sesquicentennial · 6
  - Travel Poster · 6
    - Tasmania · 3
    - Western Australia · 2
    - New South Wales · 1
  - Whaling · 5
  - Exploration · 2
  - Emigration · 2
  - Sesquicentennial, 1938 · 1
  - Great Barrier Reef · 1
  - Melbourne · 1
  - Whale · 1
  - Olympic Games · 1
  - Australian Capital Territory · 1
- Art Exhibition Poste... · 566
- Aboriginal Australia... · 315
- Music · 286
- Entertainment · 263
- War · 261
- Women's Movement · 214

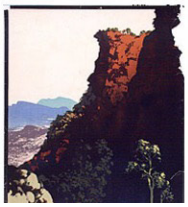
Top Artists | Media Categories | Print Types

AUSTRALIAN NATIONAL TRAVEL ASSOCIATION · 18 | Percy TROMPF · 15 | Eileen MAYO · 7 | MCLAREN · 6 | UNKNOWN · 6 | James NORTHFIELD · 6 | John BAINBRIDGE · 4 | TROEDEL &Amp; COOPER [Melbourne]. · 4 | Gert SELLHEIM · 4 | Richard BECK · 2

Decade information







5.6 Subjects Explorer

Available at:
http://printsandprintmaking.gov.au/explore/subjects-explorer/

Development

Subjects Explorer provides a unique way to explore the AP+P collection based on an artwork's subject (Figure 47). This detailed information would normally only be accessible on the individual artwork page on the main AP+P website[17]. *Subjects Explorer* aims to bring this rich information to the forefront by revealing related subjects, encouraging new ways of accessing the collection and allowing new discoveries to occur. There was a long gestation period for *Subjects Explorer*, with initial investigations into the viability of the data occurring in March 2014. Understanding the complexities of the data, and determining how to get it into an appropriate format took some time, with development of the interface itself occurring over an intensive four month period that ended in September 2015.

Following the success of utilising Angular for *The Fader* I decided to use another open-source framework, Bootstrap, in the development of *Subjects Explorer*. Bootstrap is a front-end framework that contains a number of components designed to enable faster Web development. I was particularly interested in determining how the responsive grid system, created to simplify the CSS required to ensure the interface functions effectively across multiple screen sizes, could be used in this style of interface. I will reflect further on Bootstrap in Chapter seven.

Outcome

Subjects Explorer embraces a strong column based layout informed by the Bootstrap grid system. It consists of two main sections: on the left, a hierarchical list of all the subjects from the AP+P collection; and on the right, the associated artwork information.

Figure 47 (opposite)
Subjects Explorer

17 It is important to note that not all artworks have subject data, and so, not all artworks in the collection are represented in this interface.

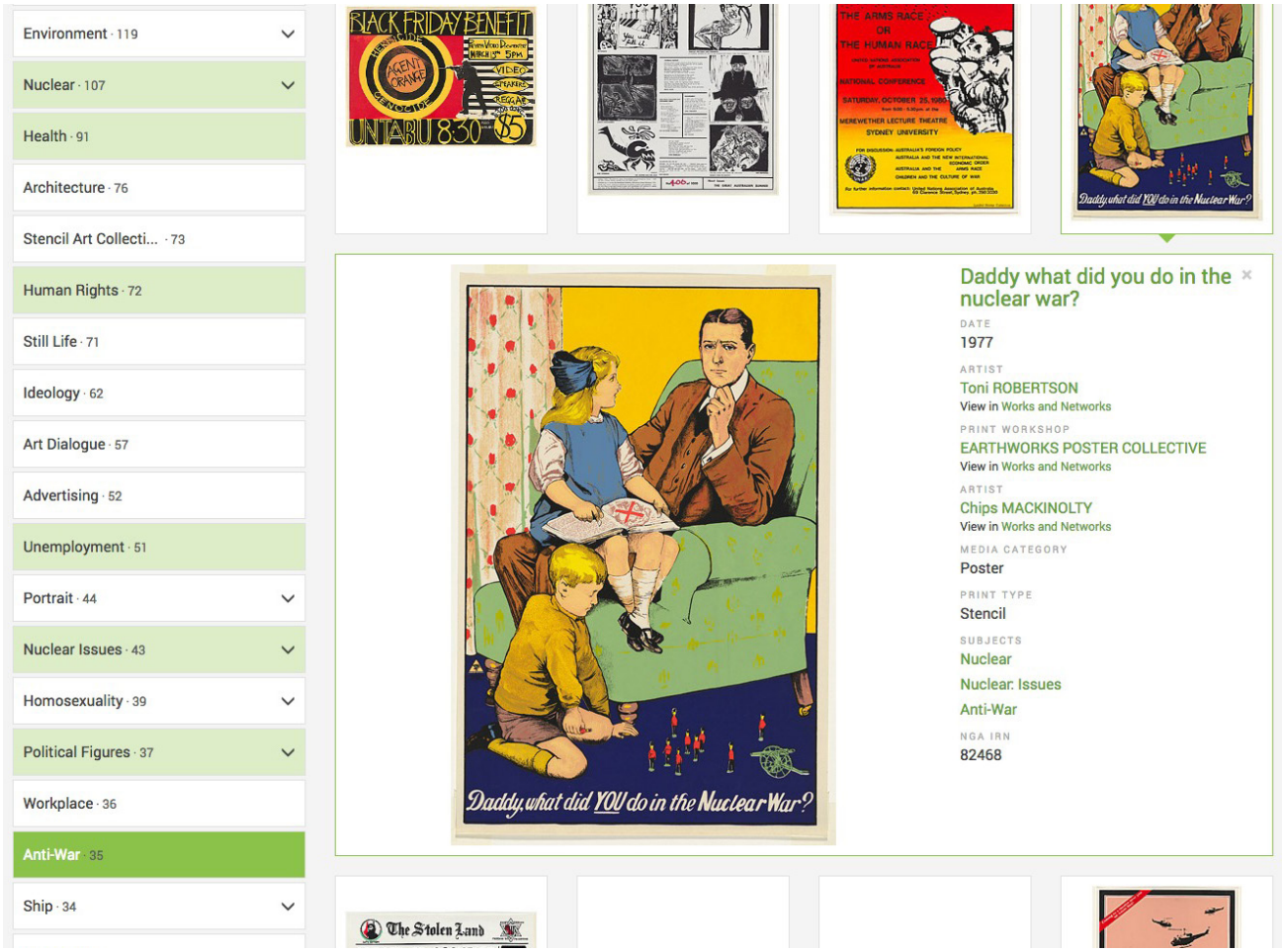


Figure 48
Detail of a selected artwork showing related subjects in left pane

The subject list displays 90 top-level subjects, many of which have nested subjects. Clicking the chevron arrow on the right of the box will reveal them—they can be up to four levels deep. Next to the subject is an artwork count where the user can quickly see that many subjects have fewer than 20 artworks. I deliberately chose to leave this 'long tail' visible, rather than grouping them together into one box (containing, for example, all subjects with less than 20 artworks) because I didn't want them to be absorbed and 'lost' within the main list. By keeping the subjects separate, a user can actually see them and explore their contents. A dark green background identifies the currently selected subject and a light green background indicates that there are related subjects within those artworks. For instance, after selecting 'Anti-War', we see that 'Unemployment', 'Human Rights' and 12 other subjects all have a light green background too; upon viewing the details for the artwork, *Daddy what did you do in the nuclear war*? by Chips Mackinolty and Toni Robertson, we can see that it has three subjects: 'Nuclear', 'Nuclear: Issues' and 'Anti-War' (Figure 48).

The section on the right takes up the majority of the screen width and contains associated information about the artworks for each subject. At the top is an interactive summary that allows the user to refine the display of the artworks below. There are three tabs which allow filtering by: top artist, media category or print type; and a bar chart to filter by decade. Under each tab there can be up to 10 buttons; in these the font size is scaled to be relative to how many items they represent—immediately giving the user a visual indication as to which button has the most records associated with it.

The artworks themselves are displayed in a four column grid with substantially larger thumbnails than in the earlier interfaces; this change in artwork image size was a key conclusion that emerged from feedback received in the evaluation study (see Chapter six). If there is no image, rather than displaying an 'x', I show the artwork title, date and an explanation 'No image available'. This technique is influenced by the Albers boxes[18] on the Cooper Hewitt site which impressed me as it offered extra information about the object, even when an image was missing. Clicking an artwork creates a focus+context display where a new full width row is inserted immediately below the selected artwork, which contains a large image and additional cataloguing information. A link to the full artwork or artist record in the AP+P site is available via the artwork title or artist name; as is an additional link to view the artist in the *Works and Networks* interface. Each subject associated with the artwork is also listed. Clicking on one of these will refocus *Subjects Explorer*, for example, changing focus from Australia to Human Rights. This allows a further method of navigating the interface without having to go through the main subjects list.

Unlike the other interfaces, *Subjects Explorer* contains brief instructions (Figure 49) about how to use the interface, which are visible when accessing it via the direct URL[19]. These are intended to resolve some of the issues of confusion as identified in the evaluation

18 Albers boxes is a technique developed by Cooper Hewitt to try and provide some understanding of an object even if the image is missing: http://labs.cooperhewitt.org/2013/albers-boxes/

19 Direct URL is: http://printsandprintmaking.gov.au/explore/subjects-explorer/

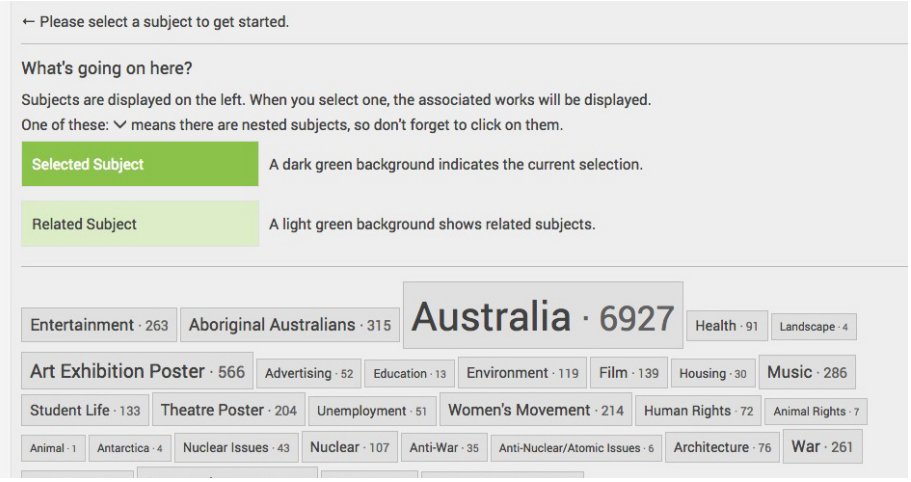


Figure 49
Instructions in *Subjects Explorer*

study. Below the instructions is a grouping of the top 30 subjects, with the font size scaled using the same technique as on the filters as described previously—the subject with the most artworks is Australia, it stands out because the font size is 41 pixels, the next largest, Art Exhibition Poster, has a font size of only 20 pixels. I call this technique visual encoding and I will discuss it further in Chapter seven. The repetition of the content provides multiple methods of accessing the data, encouraging the user to start exploring.

The interface was developed using Angular, HTML5, CSS and the Bootstrap framework. The underlying data is accessed from static JSON files that are saved on the AP+P server. I attempted to use an API to extract the data, however, unfortunately the subject data exists as a single string (even if there are multiple subjects) within the artwork record. For instance, the artwork Koala Brun [Koala, Brown] by Jacques Christophe Werner contains five different subjects and is represented in the database as: "Voyages of exploration and discovery 1600-1900: French | Australia, Animal: Koala", see Table 2.

Table 2
Subject string hierarchy

String in AP+P artwork row	**Subjects (with hierarchy)**	
Voyages of exploration and discovery 1600-1900: French | Australia, Animal: Koala	Voyages of exploration and discovery 1600-1900	Australia
	French	Animal
		Koala

Processing this data into individual subjects on the client-side for thousands of artworks would be an extremely time consuming and processor intensive task. To avoid this, I sought an alternative method of accessing the data, as I will explain. It must be noted that the final convoluted approach took some time to get right and the many steps are described in detail below.

Firstly, due to uncontrollable external factors, I was unable to get access to the MySQL database for the AP+P website until after this process was started. That being the case, I had a CSV exported directly from the NGA's KE EMu collection management system, which contained information on every single artwork in the collection. Unfortunately this file contained only the NGA Individual Record Number (IRN) rather than both the IRN and the AP+P ID (referred to as PPM ID in the data). This is because the AP+P website runs on a database external to the normal NGA collection.

After removing those artworks with no listed subject, I used OpenRefine to split the subject into new cells based on multiple identified delimiters. I knew from using similar data in my Honours project that it was extremely beneficial, at this stage of the process, to ensure the data was as clean and structured as possible. This point is emphasised by Peter Neish (2014) who explains that OpenRefine is the tool of choice for cleaning large datasets. He writes: "a highly recommended step… is cleaning up your data. This may include fixing problems such as misspellings, duplications, formatting inconsistencies or missing data." In my case, I corrected spelling mistakes in many of the subjects, and trimmed extra white space created from the splitting process. One of the most important features of OpenRefine is that it records every step taken, "so that you are able to replicate the same sequence of steps to clean up an subsequent batches of data" (Neish, 2014). Whilst cleaning this dataset, I managed to obtain another CSV file which listed the IRN for each artwork alongside the PPM ID. I used OpenRefine to

add the PPM ID from this new file to any matched artworks in my subject data. I was then able to export a JSON file directly from OpenRefine. The structure of the data at this stage was:

```
{"works" : [{
      "ppmID" : 13773,
      "ngaIRN" : 101338,
      "subjects": [
            [null,null,null],
            ["Entertainment"," Music",null,"Band",null,null,null,null]
            [null,null,null,null],
            [null,null,null],
            [null,null]
]}
```

The file size at this stage was 4.1MB—which was larger than ideal. To overcome this, I wrote a script which cleaned this file by removing the null fields and empty arrays. This resulted in a much smaller 1.2MB file, with 'clean' subject data that now looked like this:

```
{"works":[{
      "ppmID":13773,
      "ngaIRN":101338,
      "subjects":[[
            "Entertainment","Music","Band"
]]}
```

However, the data structure was 'work' based: it was a list of artworks and their subjects, rather than a list of subjects and their artworks. After the development of a Node.js script which reversed the structure and mapped it appropriately, I finally had the correct data structure:

Single subject

```
{"subject":"Entertainment",
      "items": [{
            "ppmID":13773,
            "ngaIRN":101338
            },{
            "ppmID":32087,
            "ngaIRN":84064
}]}
```

Subject with children

```
{"subject": "United Kingdom",
      "items": [{
            "ppmID":75,
            "ngaIRN":28765
      }],
      "children":[{
            "subject":"London",
            "items":[{
                  "ppmID":75,
                  "ngaIRN":28765
            }],
            "children":[{
                  "subject":"ThamesRiver",
                  "items":[{
                        "ppmID":75,
                        "ngaIRN":28765
                  }],
                  "children":[]
}]}]}
```

This data was now substantial enough to load a basic interface, however I knew that to show each artwork I would need to make a separate API call for each one, in order to retrieve artwork title, artist, date etc. This would take too long to process. Instead, I decided to harvest that data from the AP+P site using the artworks API and save it to individual JSON files for each subject.

To achieve this, I wrote a Node.js script which loaded the ppm-subjects.json file and queried the server via the artworks API to return the detailed artwork information. At the same time, I performed a number of extra functions to clean the data and format it as required for the interface to function—these functions would normally run when the data is loaded in the interface. The resulting data for each subject was then saved to the server.

The new data structure was:

```
{
      "subject":"Nuclear",
      "date":0,
      "title":"NoBase,noNucleartarget",
      "cleanSubject":"nuclear",
      "subjectPath":"/subject/nuclear",
      "ppmID":27718,
      "ngaIRN":179534,
      "hasImg":true,
      "imgURLMed":"http://artsearch.nga.gov.au/
            IMAGES/MED/179534.jpg",
      "mediacat":"poster",
      "print_type":"stencil",
      "artists":[{
            "name":"PRAXISPOSTERWORKSHOP",
            "role":"artist",
            "id":13557
      }]
}
```

Finally, when the interface is loaded, I used the 1.2MB 'clean' subjects JSON to generate the subjects listed in the left pane. When a subject is clicked, the individual subject file is loaded. As this JSON file already contains all the artwork data, the interface loads extremely quickly with the only delay being caused by images downloading from the NGA server (see Figure 50 for an illustration of where the data is loaded from for the different components of the interface).

The final technique of processing the data was complicated and would be difficult and time consuming to reproduce, however it did highlight a number of data-related issues common with this style of artwork which I will reflect on in Chapter seven.

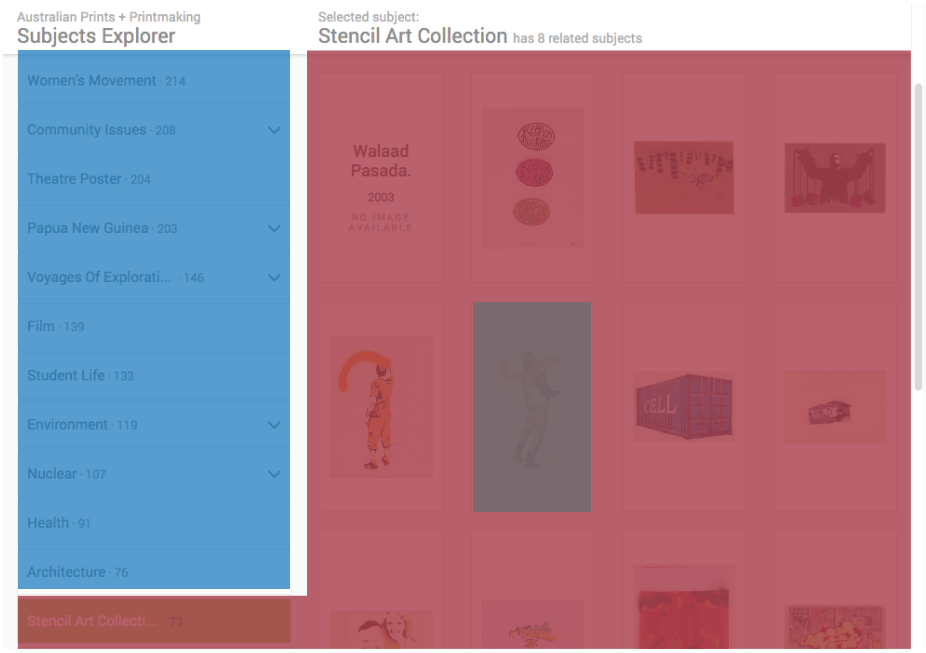


Figure 50
Subjects Explorer data loading process

Australian Prints + Printmaking
Artist Timeline

Timeline focus:
Robin White 1946 - Female

121 works 7 exhibitions 10 references 6 galleries

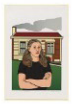
1946

Born July 12, 1946
in Te Puke, Bay of Plenty, Aotearoa New Zealand

1972

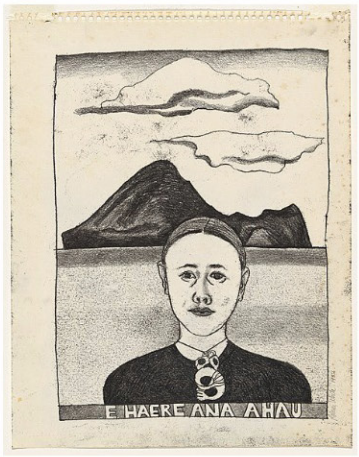
3 biennale internationale de l'estampe.
AT
Musee Galliera.

1979



1982

E haere ana ahau (I am leaving)
DATE
1982
ARTIST
Robin White
View in Works and Networks
MEDIA CATEGORY
Print
PRINT TYPE
Monoprint
NGA IRN
10145



1983

(Title page)
No image





1985

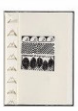





Not titled (Brigid is holding a nautilus).
DATE
1985
ARTIST
Robin White
View in Works and Networks
BOOK-BINDER
BOYD BOOKBINDING CO LTD.
View in Timeline
View in Works and Networks
MEDIA CATEGORY
Artist's Book
PRINT TYPE
Relief
NGA IRN
163667








5.7 Timeline

Available at:
http://printsandprintmaking.gov.au/explore/timeline/

Development

Timeline provides the user with an immediate visual overview of the life of an artist or gallery (Figure 51). For the first time, it was possible to see detailed information about artworks, exhibitions and references all within a single view. Development of *Timeline*, my last interface for the AP+P collection, took three months and it was released online in October 2015.

Whilst *Timeline* was the last interface produced, the idea for it was actually one of the first I had when conceptualising the research project. I knew from the outset that the data contained within the AP+P collection was very detailed, and that it went beyond the traditional level of cataloguing information available online (which is why it was developed as a stand-alone website). The data includes well structured connections between artists and artworks; artworks and exhibitions; exhibitions and galleries; artists and galleries; artists and references and so on. As a result, I envisioned an interface that combined all of this rich data into one engaging display which I was certain would be extremely useful to the user and an interesting one to explore.

I experimented with the timeline approach in 2012 (using HTML, CSS and jQuery) and in 2013 (using a popular JavaScript library TimelineJS). These experiments showed that while the approach had potential, there were many issues to resolve. For example, in the 2013 attempt, there were mass groupings of artworks overlapping each other at the beginning of each year with long gaps between them, caused by the artwork creation date lacking a day and month. Consequently, I left the development of *Timeline* until last because I did not actually know if it would be possible to build such an interface.

Figure 51 (left)
Timeline for Robin White

Following on from the success of *Subjects Explorer* I decided that creating *Timeline* would be the best way to bring together many of the concepts formulated over the development of the previous five interfaces. I aimed to, once again, investigate ways of providing an overview (for different types of data, within the same display), as well as experimenting with new 'focus+context' techniques.

Outcome

Upon accessing *Timeline* directly, it defaults to an artist centric view, where the user is presented with 70 random boxes containing different artist names and artwork count. The font size in each of these is changed using the same technique as in *Subjects Explorer*. This provides an immediate visual clue as to which artists might have a more developed timeline. Above these boxes is a search field, which allows a quick search for other artists represented in the collection. I added this function to make the interface as useful as possible, as it enables those users who already know the collection or are looking for a particular artist, to easily find them and view their timeline. Even though the initial view of artists' names is randomised, the order is retained so that if a user returns to the start screen by using the back button, they will still be able to see the original selection of artists. Clicking a name will load the *Timeline* for that artist.

Timeline features a three-column display. To the left, exhibitions, galleries and references are displayed; the middle pane contains the year; and on the right are the artworks themselves.

The artworks, exhibitions, galleries and references are displayed in boxes, each of which has a colour coded left border. These colours align with a label and count in the page header, with light green for exhibitions, blue for artworks, purple for references and dark green for galleries. This technique is also used in *All Artists* and allows a user to understand what type of data is represented in the different boxes on screen. The page header, which also contains the current artist's name, stays at the top of the window as the user scrolls down, subtly alerting them to the current focus and what the coloured borders represent.

The left pane contains boxes with information on exhibitions and references. For an exhibition, I show the title, primary gallery name and total of artworks included in the exhibition (if this data is available); a reference box contains the title and author. All are interactive and clicking on a box will either: expand the current box to show more information, or for an exhibition, insert a new full width row into the interface (under the current year) containing detailed exhibition information. This method of adding extra information to the interface is part of a visual language developed over the creation of the previous *Explore* interfaces.

The year labels are shown in the middle pane, joined by a grey-blue coloured line that extends between each one. Where there is no data for a particular year, the line is broken and the label removed, thus significantly reducing the page height and ensuring as concise a display as possible. For example, Margaret Preston was born in 1875 and her first artwork was 40 years later,[20] each label is 30 pixels high which would have added 1200 pixels of blank space to the start of the interface.

Unlike most timelines, this one often continues after the artist's death (and into the future), and so I have designed it to accommodate exhibitions and references created past the current year, ensuring longevity of the interface. At the bottom are the records that have an unknown date, these are grouped together and are always available; removing them from the interface would skew the representation.

The right pane contains information about artworks produced in each year. These are displayed in a box that either contains a thumbnail image of the artwork, or the title and text: 'No image'. The boxes are interactive and clicking on one will cause it to expand and show a large artwork image alongside the title, date, list of creators, media category and print type (Figure 52). Additional links to *Timeline* and *Works and Networks* are included under each artist's name, in order to encourage the user to view the artist in one of the other *Explore* interfaces.

20 This is the first artwork recorded within the AP+P data, she may have produced other earlier artworks not held in this collection.

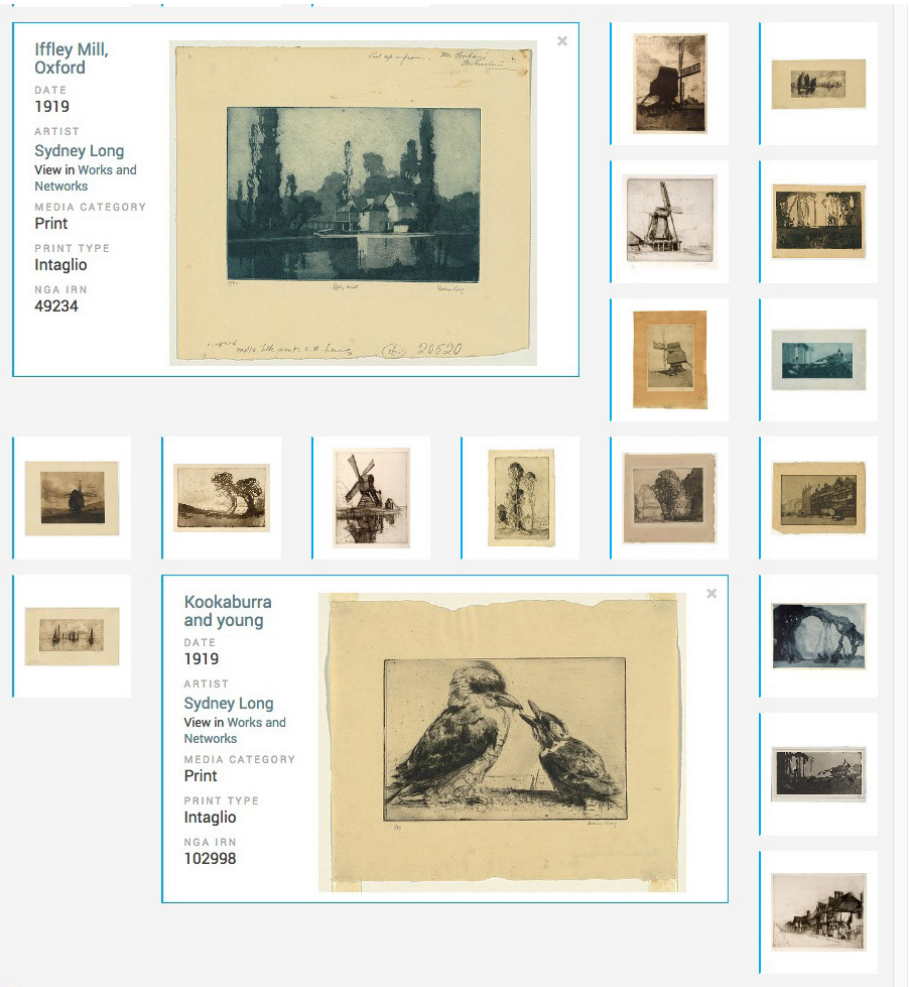


Figure 52
Detail of expanded artwork amongst thumbnails of other artworks

The detailed exhibition view (Figure 53) is one of the most exciting features of this interface as it shows how richly structured data can be used to create engaging new methods to represent collection data. Many of the exhibitions catalogued in the AP+P collection also contain information on which artworks were exhibited. This interface presents these artworks in both the context of an artist's life span and amongst artworks produced by fellow artists.

The expanded exhibition row maintains the exhibition border colour at the top and bottom and has a white background to clearly differentiate it from the rest of *Timeline*—thus allowing me to introduce an additional layout for the information within the box. Two panes are in this extra display, the left contains detailed exhibition information, and the right, the artwork data.

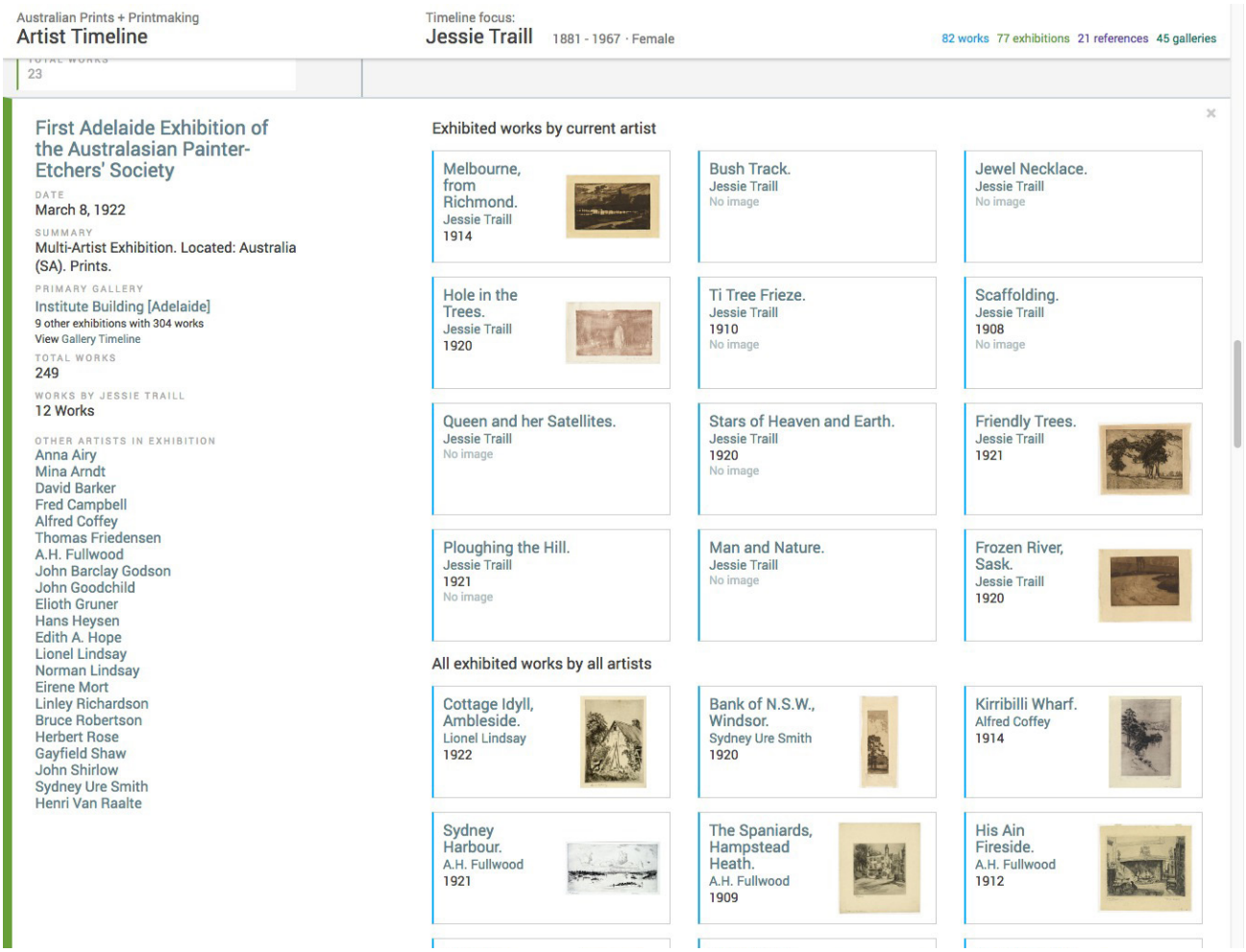


Figure 53
Detail of expanded exhibition view in artist *Timeline*

I have made the representation of data within this focus+context display as rich as possible. Therefore, in addition to detailed exhibition information containing the title, date and summary, I have also included the name of the primary gallery, how many other exhibitions were held at the gallery and the total number of artworks in all of their exhibitions; as well as the total count of artworks in the current exhibition, how many artworks were by the current artist and a list of other artists in the exhibition. The timeline can be refocused from within this pane by clicking on an artist's name or following the 'View Gallery *Timeline*' link below the gallery name. I call this technique a 'dynamic focus+context display' and I will discuss it more in Chapter seven.

The right pane contains a grid of artworks from the exhibition. Exhibited artworks by the current artist are shown first (if that information is in the AP+P collection); followed by

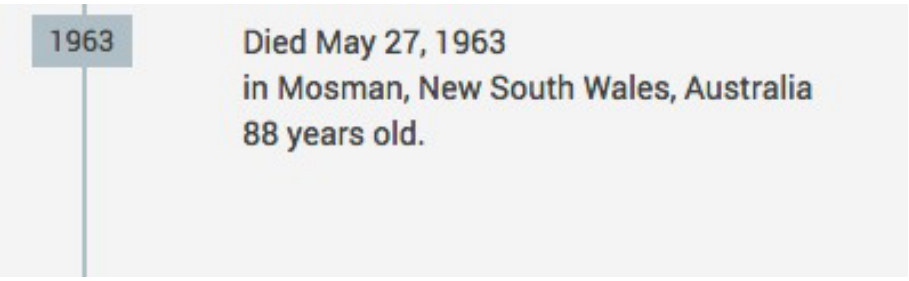


Figure 54
Detail of data-driven descriptive text in *Timeline*

exhibited artworks by all other artists below. For each of these, rather than just showing the thumbnail of the artwork, I include the title, primary artist name and artwork date. A link on the title will take the user to the original cataloguing record on the main AP+P site. Clicking the artist name will refocus the timeline. These artworks are sorted to display those with images first, as this indicates that they are probably in the NGA collection.[21]

Where possible, I have added additional data-driven descriptive text to the timeline to help reveal further aspects about the artist (Figure 54). For instance, in the first year of the timeline, the exact date of the artist's birth and their birthplace is listed; if the data contains a death date and place, then it is listed beside the corresponding year also—if the data contains both the birth and death years then I can calculate the age of the artist when they died, and that is listed too. I have found this simple feature creates a more humanising interface. This style of turning simple catalogue data into more meaningful information has also been demonstrated on the Cooper Hewitt collections site[22], which I will discuss further in Chapter seven.

I created the interface in a way that allowed me to generate a timeline for an artist or gallery, without requiring a completely new design or rewriting of the code; many of the functions are shared between both views. This flexibility means that regardless of the focus, the interface is able to adapt to differing data sets accordingly.

21 The AP+P site includes records of artworks that aren't necessarily in the NGA collection. It was intended as a record of all printed material created in Australia.

22 See for example: https://collection.cooperhewitt.org/objects/18643215/

For the Gallery *Timeline*, exhibitions and references are shown in the left pane. On the right, the artworks within each exhibition are grouped together, these are always shown—in contrast to the artist *Timeline* where a user needs to 'open' the exhibition. Because there is much less data to load in the gallery *Timeline*, I could show all the artworks without it taxing the browser or server.

As with *Subjects Explorer*, *Timeline* was built with Angular, the Bootstrap front-end framework, HTML5 and CSS. In regards to the data, I wrote a number of custom API calls that return data from the various entities: artists, artworks, exhibitions, galleries and references. In order to reduce the loading time of some components of the interface I created a PHP script which saves an API call (to return a list of the artists) to a static JSON file on the server. This script runs on the server every 24 hours in order to ensure the data remains up to date. The load time from a static file is considerably faster than waiting for the API call to be returned asynchronously from the server, additionally, it ensures there are no undue processes being run on the server.

5.8 Conclusion

In this chapter I have described the conception and development of the six interfaces produced as part of my research project. I have also outlined how I moved from working collaboratively with Whitelaw to independently building increasingly complex interfaces. The unique nature of the creative works represents a contribution to knowledge in themselves. I have demonstrated how each interface offers users distinct ways to access the AP+P collection by building upon the concepts outlined in my literature and practice reviews, and introducing new methods and techniques for providing access to cultural heritage data. The most rewarding of these techniques has been the creation of comprehensive focus+context displays, which provide access to fine-grained detail whilst maintaining context: these are in marked contrast to traditional collection access interfaces.

I have shown how this technique can be applied in many different forms, from the three-pane display in *Works and Networks* to the expanded exhibition view that is inserted into *Timeline*. It is a technique which is particularly effective when combined with an overview, as illustrated in *Decade Summary*. The interfaces demonstrate the strength of combining data visualisation techniques—particularly size, scale and colour—with contemporary Web design conventions, in order to create novel representations of the collection.

The next chapter will discuss the evaluation process and the results of a mixed-method evaluation study which sought to evaluate *Works and Networks*, *Decade Summary* and *All Artists*. In Chapter seven, I will outline my contributions to new knowledge and reflect on the practice-based works through a discussion that addresses my research questions in depth. Finally, in Chapter eight, I will summarise the project, its contributions to knowledge and discuss how the techniques I have developed are applicable to any digital collection.

6 Evaluation

6.1 Introduction

In the previous chapter I outlined the six creative works produced as part of my research project and detailed the techniques I developed to reveal different aspects of the AP+P collection and encourage exploration and discovery. In order to examine the techniques' effectiveness, I completed a two stage evaluation process which is the subject of this chapter.

The first stage consisted of a mixed-method evaluation study, which sought to evaluate the first three interfaces, *Works and Networks*, *Decade Summary* and *All Artists*. The completion of this evaluation study was a formative component in my research project, as the results directly informed the development of *Subjects Explorer* and *Timeline*. The study consisted of four think-aloud observations made by selected participants, an open Web survey and associated data logging of usage. In the discussion that follows I outline how, through the collection of quantitative and qualitative data, I was able to effectively address the aims of the evaluation study. I will then introduce each method and describe how the data was collected and analysed. Following this, I discuss the results starting by introducing the coding terms and providing a summary of the data collected, and then examining the results for each method in detail. The findings of the study's will show how the interfaces succeed in encouraging exploration and discovery. Finally, I reflect on the mixed-method evaluation study as a whole and identify the key challenges faced in subsequent work.

The second section of the chapter considers the final two works produced, *Subjects Explorer* and *Timeline*, which attracted critical feedback and evaluation via presentations at numerous conferences and workshops in Australia.

6.2 Mixed-method evaluation study

The purpose of the evaluation study was to assess the effectiveness of *Works and Networks*, *Decade Summary* and *All Artists* in encouraging exploration and discovery. The structure of the chapter is as follows: I will outline the evaluation methodology and the research questions posed. I will then discuss the three different methods used in the study: think-aloud observations of selected participants, a Web survey and data logging. Following this, I provide a summary of the data collected, and a detailed discussion of the results. Finally, I will identify a number of key conclusions of the evaluation study that informed development of *Subjects Explorer* and *Timeline*.

As discussed in Chapter four, Pohl et al. (2010) suggest that the exploratory nature of these new types of interfaces can make the evaluation process difficult, resulting in the need for alternative models of evaluation. Driscoll et al. (2007) explain that a mixed-method approach refers to the collection of both qualitative and quantitative data in the context of a single study. In a similar Web-based study of visual exploration tools, Dörk and his collaborators (2012) utilised a mixed-method approach. They argue that this format provides more information about both human use and system performance, which, as visual exploration is a relatively new field, will lead to a better understanding of the exploration process and the challenges it poses (Dörk et al, 2012). There were clear parallels between Dörk's exploration-focused project and my own, and as a result, I adopted a similar mixed-method evaluation study format.

The components of the study – think-aloud observations, a Web survey and associated data logging – were selected due to their complementary nature and the anticipated high quality of the results they would provide. It was through the collection of both qualitative data (through the think-aloud observations and part of the Web survey) and quantitative data (from the Web survey and data logging) that I expected to gain an in-depth understanding of how the interfaces are used.

I sought to answer a number of specific questions through the completion of the evaluation study. They were:

> *Was the user able to engage in open-ended exploration and discovery?*
>
> *What particular aspects of the interfaces did the user engage with?*
>
> *Is visual exploration a useful technique for encouraging discovery?*
>
> *How do people use these interfaces?*

The evaluation study was approved by the University of Canberra's Human Research Ethics Committee.

Think-aloud observations

The first method consisted of think-aloud observations conducted with four participants. North (2006) calls for researchers to embrace an open-ended protocol where the user is able to interact with the data in a way they choose, specifically, by thinking aloud. While the participant verbalises their thoughts the researcher observes and records their comments (Holzinger, 2005).

Lewis & Rieman (1994) remark that although the think-aloud method is very simple, the comments provide a rich trove of information. Zahran et al. (2014) explain that a small number of users completing a think-aloud observation can result in a wealth of qualitative data. The observational approach is a means of gaining a direct insight into the knowledge and methods of human problem-solving (van Someren et al., 1994) and provides context for and interpretation of the participant's activities (Pohl, 2012).

Holzinger (2005) notes that the time needed to analyse a recorded think-aloud observation is substantial. In order to gain thorough insights, without being overwhelmed by data, I decided to conduct four think-aloud observations.

Methods

Participants for the think-aloud observation were carefully selected to represent a wide sample of potential users of the AP+P website. Of the four participants, two came from within the arts community in Canberra: one as a printmaker and the other as a staff member in the Department of Australian Prints and Drawings at the NGA. These two participants were experts in their field of knowledge, they knew the collection and how it was structured. The other two participants were unfamiliar with the NGA print collection. One was a scientist, the other a second year economics student. The varied backgrounds of the participants were extremely important as they allowed me to consider how the research questions applied to different users in different contexts. For example, one of the research questions asks "*is visual exploration a useful technique for encouraging discovery?*"; the answer to this question may depend on the background of the user interacting with the interface. If the user already understands the collection they might approach the idea of visual exploration very differently to someone who has no understanding of the collection. The participants were of varying ages, from the mid twenties to early fifties, and represented an even gender split. Their computer literacy levels ranged from proficient to expert.

The observations were conducted at locations chosen by participants, with all but one being at their workplace. They were provided with a voucher of monetary value as a form of compensation for their time.

Each observation followed the same format: I introduced the AP+P website to the participant using deliberately informal language and asked them to 'play' with it, starting with 'Search' and then moving on to the *Explore* section. There was no defined goal and no time limit. As the participant interacted with the AP+P website, a program called Silverback[1] recorded all screen activity as well as anything that was verbalised.

1 Silverback: https://silverbackapp.com

Analysis

Yang (2003) explains that data analysis of verbal protocols is typically a three step process: firstly, the recordings are transcribed; then they are segmented; and finally they are encoded. Nielsen et al. (2002) elaborate on these steps noting that while the transcription and segmentation are important, the main part of the analysis is the "actual encoding of the verbal reports."

Ericsson and Simon (1993) argue that the researcher should make minimal use of context (for example, being aware of the interface) when analysing the data in order to avoid any possible bias that they may introduce during the segmenting and encoding processes. This is an interesting observation but is at odds with my approach where context (the interfaces) is of prime importance. My reading of the literature reveals that most verbal data is transcribed from an audiotape or recording (Ericsson and Simon, 1993; Yang, 2003). In my study I used both video and audio recordings that required analysis concurrently. The video is integral to understanding the context of what think-aloud participants are describing. Often behaviour that occurs in the video also needs to be transcribed, for example, if a participant clicks a particular link within the site or moves between different interfaces. Therefore, in contrast to the view raised by Ericsson and Simon, I had to maintain the context of the video whilst analysing the data.

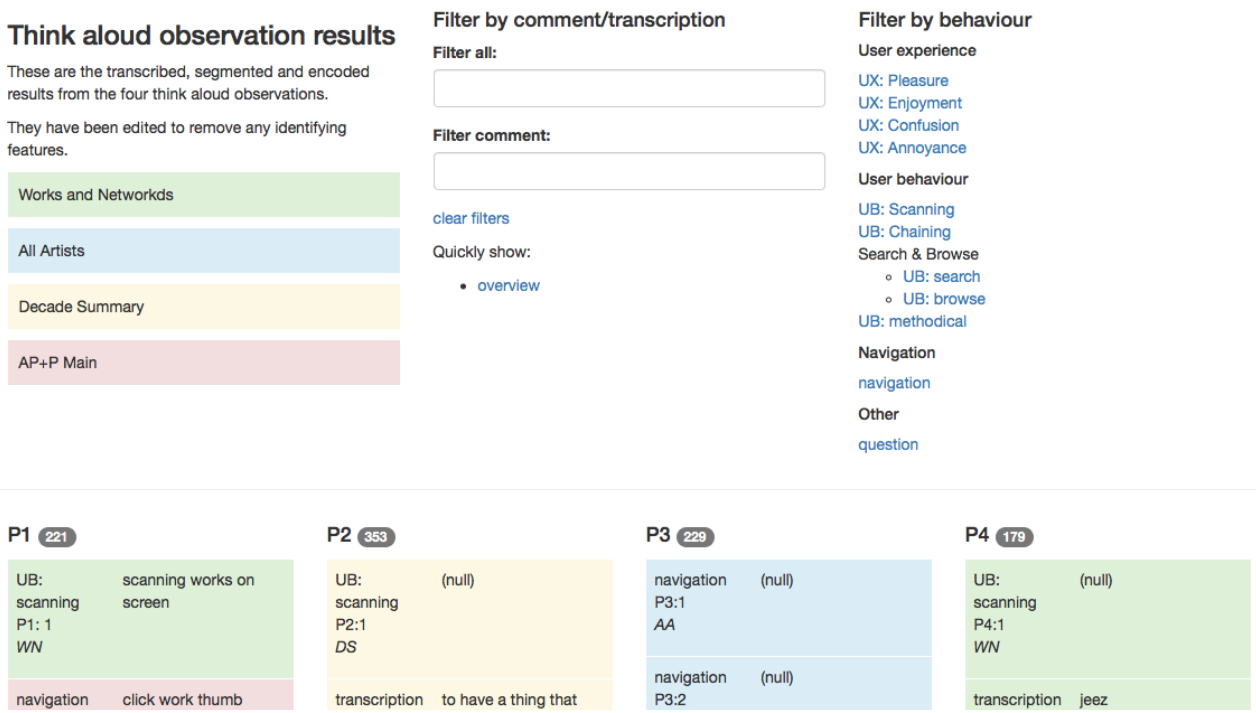
Goransson et al. (2007) explain that "any researcher who has completed protocol analysis by hand knows that the sheer amount and complexity of the process can be overwhelming." To determine how to work most effectively with the data I conducted a survey of the various digital qualitative data analysis tools that are available. There is a vast number (Bryman, 2008) ranging from popular stand alone programs developed by large companies, for example, QSR NVivo, MaxQDA and Atlas.ti; to open source and independently produced software such as Time-ART, ELAN and VCode & VData. After analysing a number of them I decided that rather than using a complex proprietary-based system like NVivo, independent software was preferable because of its flexibility.

I conducted an in-depth examination of four tools: ELAN; VCode & VData; Annotating Academic Video (AAV) and the Open Video Annotation (OVA) project. Through trialling each program I discovered that both OVA and AAV were not stable enough to be used and that ELAN is best suited to complex linguistic analysis. As a result, I selected VCode & VData (Hailpern and Hagedorn, 2008) as the most appropriate software for the type of qualitative data analysis I wanted to undertake. Whilst the program is relatively old (from 2008) it was the most stable and allowed simple video annotation to be combined with identifiable events.

Before starting the segmentation and encoding process I viewed the videos and removed unrelated information, for instance, the initial recorded introduction to the AP+P website. Whilst doing this I made a list of the key behaviours that participants displayed and then used them to configure the coding template.

In some research studies, the process of transcription, segmentation and encoding are completed in discrete steps, however, because of the nature of the data and its complexity, I chose to combine them. This approach was also taken by Norgaard and Hornbaek (2006) who applied keywords to each segment as they

Figure 55
Think-aloud component of evaluation website



analysed their data. It is an iterative process that can be subject to "several cycles of interpreting, defining and refining" (Yang, 2003). I found that my participants regularly engaged in multiple behaviours at once, which resulted in a number of overlapping behaviours.

In order to refine and encode the data I followed a bottom-up analytic approach, where the aim is to find relationships and patterns in the data (Wirsch, 2014). This entailed:

> Reading all feedback provided to gain an overview of all the responses;
>
> Segmenting the different feedback into coded comments based on whether they were positive, negative, providing general feedback or suggestions;
>
> Refining these into common topics based around the particular interfaces or themes that emerged.

In contrast to other qualitative data analysis tools, VCode is designed to only perform the task of annotating and encoding the video. Once the video has been encoded the raw data can be exported and used for analysis externally. To facilitate this, I developed a Web application[2] that presents an overview of all the data, which is filterable by a query box or a set of predefined filters (Figure 55). The display is split into two sections: the filtering component at the top and the data below. The data is displayed in four side-by-side columns, one for each participant. Each column includes all of the segmented data for that participant, separated by a horizontal line. Each segment includes the coded word and the transcribed data.

The predefined filters are the same behaviours that were used in the initial coding process in VCode. They enable the user to immediately refine the data displayed on screen to only include the selected behaviour. For instance, if a user were to click the 'user experience: confusion' behaviour, the display immediately changes to only include the five segments that have that code attached. There

2 The evaluation Web application is available here: http://phd.beneb.com/evaluation-analysis

is also the ability to refine the display by filtering the transcriptions themselves, for example, by typing the word 'overview' into the query box, the user can see that Participant 3 said the word three times.

For this particular set of data, I found that the best approach was to use the inbuilt browser 'find' function alongside the various filtering techniques and combine them with manual processes like reading the actual transcriptions and writing down the key ideas.

Web survey

The Web survey formed the second component of the mixed-method evaluation study. It was designed to collect generalised and unstructured qualitative data along with basic quantitative data from a wide range of participants.

Web surveys have a number of benefits. They are convenient for respondents as they can answer the survey at their own pace, whenever and from wherever they choose (Callegaro et al., 2012); and due to the globalised nature of the Internet, they are low cost and give easy access to a broad scope of potential respondents (Hewson, 2008). The format of the Web-based survey makes it easy to understand (Driscoll et al., 2007), as it presents a user-friendly interface, which allows users to view the questions, input their answers and easily submit the form (Hewson, 2008). A Web survey containing open-ended response fields allows participants to post extensive comments (Driscoll et al., 2007).

There are limitations in completing a Web survey, as respondents are required to self-select (Zhang, 2000), which means that it can be difficult to gain a representative population sample or introduce sample bias (Callegaro et al. 2012; Hewson, 2008). However, this was not a concern in my study, as the purpose of the research was not to collect a representative sample of all users but instead to seek generalised feedback from the relatively small number willing to complete the Web survey.

In their paper, *Museum Web search behavior of special interest visitors*, Mette Skov and Peter Ingwersen (2014) describe the use

of a Web questionnaire survey consisting of both closed and open-ended questions, in order to collect information about users' "areas of interests, purposes of visiting the museum Website, preferred data elements, as well as demographic data." Additionally, Paul Marty (2007; 2011) has conducted Web surveys to collect information about the motivations of museum website users. These studies show the appropriateness of using this method in my mixed-method evaluation study.

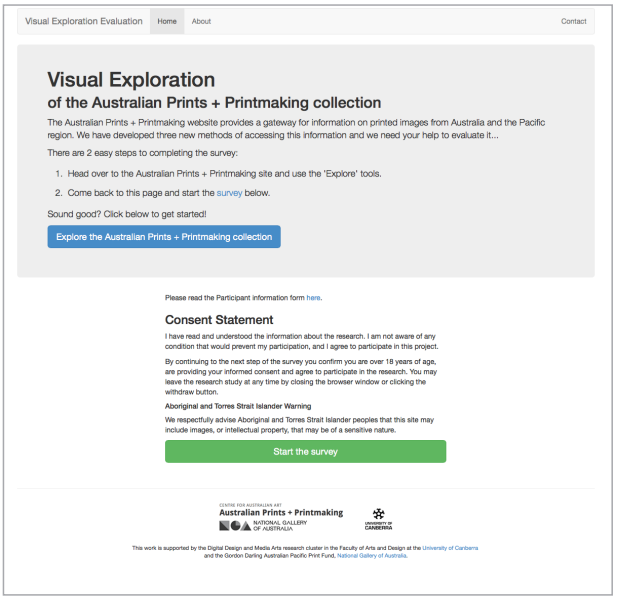
Methods

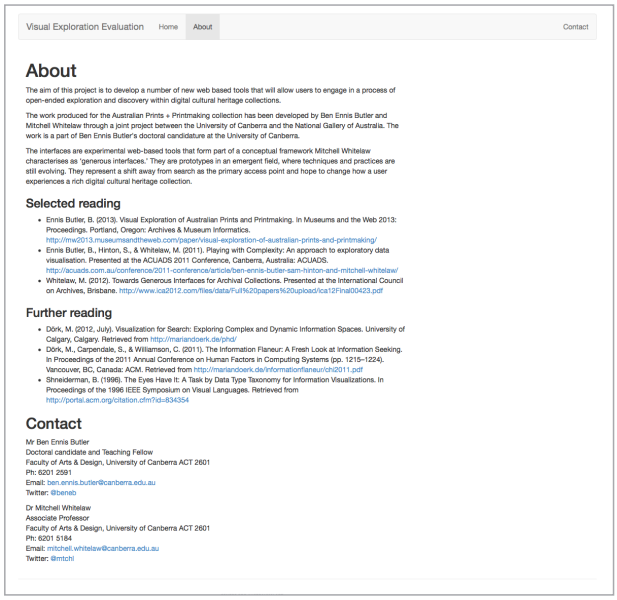
The Web survey was accessible as a stand alone website entitled 'Visual Exploration Evaluation' (VEE) with its own URL: http://visualexplorationevaluation.com[3]. Access to the survey was open to anyone, however the experimental nature of the interfaces restricted their functionality to only working on modern Web browsers Mozilla Firefox, Apple Safari and Google Chrome. If the potential participant was using Microsoft Internet Explorer they received an error message and were prompted to use an alternative browser.

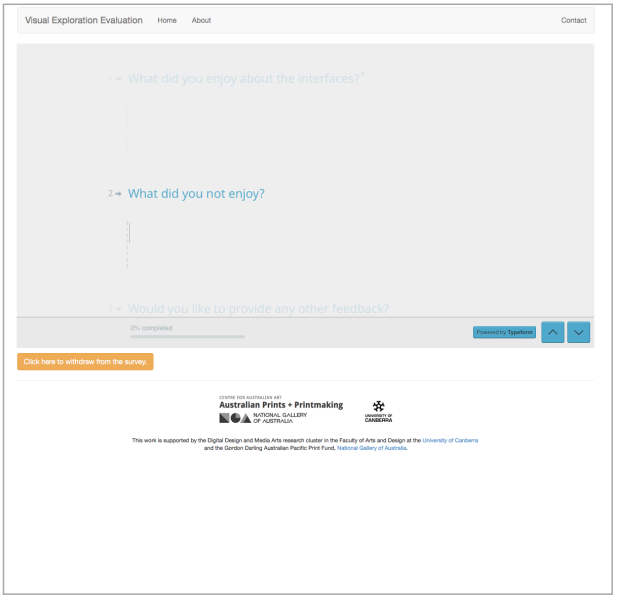
Participants were recruited to complete the survey through invitation links added to the *Works and Networks*, *Decade Summary* and *All Artists* interfaces. The survey was also publicised through the main AP+P website, the AP+P Facebook page, on Twitter, through word-of-mouth and via email.

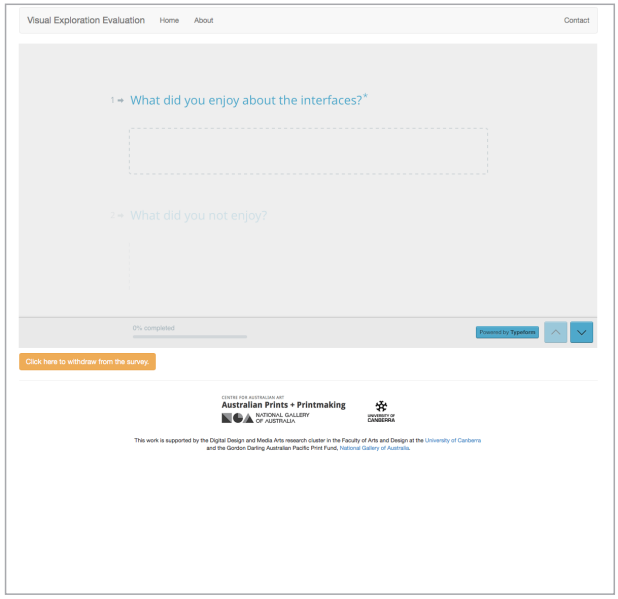
Once they accessed the VEE website potential participants were asked to go to the AP+P website (Figure 56) and use the *Explore* tools; then return to the VEE website and start the survey (Figures 58, 59). After clicking the 'Explore' button, the AP+P website would open in a new tab, allowing the potential participant to return to the VEE website at any time.

3 Accessible at http://phd.beneb.com/vee-website









The survey was created using *Typeform*[4] and was broken into two sections with seven questions in total. In the first section generalised qualitative data was collected by asking the participant three questions:

What did you enjoy about the interfaces?

What did you not enjoy?

Would you like to provide any other feedback?

Screenshots of VEE Website, clockwise from top left

Figure 56
VEE website: welcome page

Figure 57
VEE website: about page

Figure 58
VEE website: start of survey

Figure 59 (bottom right)
VEE website: survey interface

4 A web-based survey company with a focus on creating 'beautiful' forms: http://typeform.com

Underneath each question was an open text field in which they could respond. The questions were open-ended and, as with the think-aloud observations, used deliberately informal language.

The second section of the Web survey included a number of optional questions to collect basic demographic data on the participant. The questions asked were:

What is your age range? There were four options: 18-24; 25-44; 45-64; 65+

What is your gender?

How would you rate your level of experience with the internet? With a star scale, 1 being a beginner and 5 being an expert.

Would you like to receive a copy of the research results? If participants answered yes to this question, they were prompted to enter their email address.

Analysis

In order to analyse the data I segmented the comments each participant provided and encoded them according to the various points that were being discussed. For example, some participants made clear references to aspects of a particular interface they enjoyed, whereas others only made general comments about the work as a whole.

As the questions in the first section of the survey explicitly asked what a participant did or did not enjoy, they already led to some predefined encoding through positive (did enjoy) and negative (did not enjoy) comments. Additionally, as the analysis of the Web survey data progressed, a common set of terms emerged.

As with the think-aloud results, I created a Web application[5] that enabled me to easily view and analyse the data collected in the Web survey (Figure 60). The display and data handling functions are much the same, with the user able to either view all of the data or filter it using a query box. The data for each participant is displayed

5 Accessible here: http://phd.beneb.com/evaluation-analysis

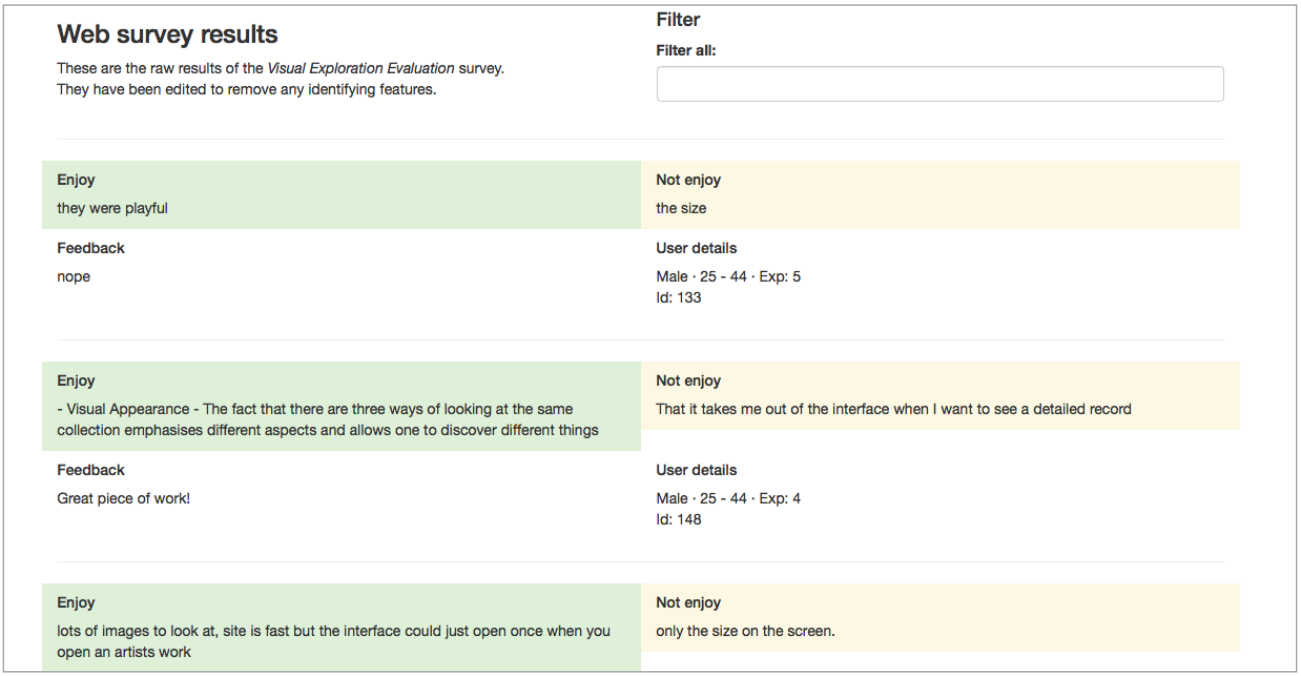


Figure 60
Think-aloud component of evaluation website

as a block and separated by a horizontal line. Each block is split into two columns: the left column containing aspects they enjoyed and those that they did not enjoy are on the right— colour coding provides contrast and allow quick identification. The block also includes any other feedback and the basic user details captured.

Data logging

The third component of the mixed-method evaluation study involved automatic data logging of usage on the *Explore* section of the AP+P website. This form of data gathering is known as Web analytics (Fang, 2007) and is defined as:

> the objective tracking, collection, measurement, reporting and analysis of quantitative Internet data to optimise websites and Web marketing initiatives. (Kaushik, 2009)

In other words, the collection of Website usage data which, when analysed, allows for insights and understandings into how a website (or interface) is used. Traditionally the use of Web analytics has focused on measuring the effectiveness of online marketing and conversion rates (e.g. into a sale). Interestingly, Clark et al. (2014) explain that this narrow view of online usage may limit the relevance and effectiveness of the analytics service when considering other kinds of online usage.

The quantity of data captured through the analytics service is much greater than could be collected through the think-aloud observations or Web surveys, as the analytics server automatically logs detailed non-identifiable user information, including data about browser type, screen size, operating system, location and so on. This quantitative data provides further insights into usage patterns and interactions that complement the data collected in the other two methods.

The report *Let's Get Real, How to Evaluate Online Success?* commissioned by Culture24 (a UK cultural group) and written by Jane Finnis, Seb Chan and Rachel Clements (2011), recommended the adoption of standardised analytics tools (specifically Google Analytics) across gallery and museum websites. The use of analytics data is also encouraged by Steven Turner (2010) in his paper *Website statistics 2.0: Using Google Analytics to Measure Library Website Effectiveness*.

Methods

Many companies offer Web analytics tools including: Yahoo Web Analytics, Piwik, StatCounter, Alexa and Google Analytics (GA). GA was already in use on the main section of the AP+P website and so it was also implemented on the *Explore* section.

On a traditional website every time a user navigates to a new page, an individual URL is loaded. For example:

Homepage (index.html) > Search (search.html) > Results (results.html)

In these circumstances, usage data is captured automatically by tracking code which is inserted into every page. However, as explained in chapter 4, the explore interfaces load content dynamically within the same HTML document—thus not actually loading a new page URL. This creates a problem because the standard tracking code does not understand that interaction might be occurring within the page. To overcome this, special GA event actions were added to the JavaScript code that generates the interfaces.

This enabled GA to capture further detailed information about the interfaces and how they were used.

The data collected by GA is grouped into reports that cover three main categories: Audience information (user location, type of browser, operating system); Acquisition (how the user got to the site); and Behaviour (what they did on the site, what pages were visited and events triggered).

Analysis

Analytics data for the *Explore* interfaces has been captured since their launch in April 2013, however, to ensure consistency of results between the various methods of the evaluation study, the analysis and discussion of the results only includes data gathered between December 29, 2013 and July 1 2014—the period when the Web survey was operational.

Data collected via Google Analytics is accessible through Google's own Web-based interface. This interface contains a number of features designed to allow a user to view different reports about the data, including fine-grained analysis of what has been collected. Additionally, the raw data is accessible through the interface or by exporting it into CSV format.

A detailed discussion of results from the data logging will follow later in this chapter.

Summary of data collected

Think-aloud observations

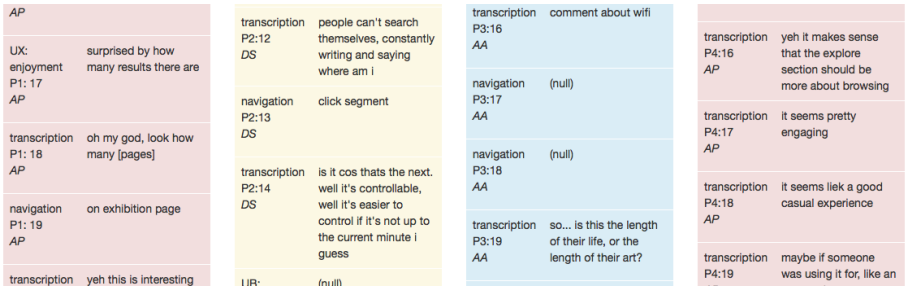
Data from the think-aloud observations was segmented and encoded, resulting in 17 different coding terms and 982 segments. The breakdown of these across the four observations was:

Table 3
Think-aloud data breakdown

Participant	Segments	Length of observation
P1	221	23 minutes
P2	353	37 minutes
P3	229	25 minutes
P4	179	19 minutes

Each segment included the coding term or a transcription of the voiced observations—or sometimes both (Figure 61).

Figure 61
Example of segmented data



The behaviours consisted of four different categories: user experience, browser behaviour, navigation and interface. Each of these contained a number of sub items which were used in the encoding process. They were:

Table 4
Think-aloud behaviours

Categories	Sub items
User behaviour	scanning, chaining, search and browse, methodical
User experience	enjoyment, surprise or delight, confusion, annoyance
Navigation	navigation through the site
Interface	Decade Summary, All Artists, Works and Networks

The summary that follows will provide a background of the coding terms and their representation within the collected data.

User behaviour (umbrella term)

This is the umbrella term used to categorise the methods of interaction behaviour the participant engaged in when using the interfaces. The term includes four sub items: scanning, search and browse, methodical and chaining. There were 78 occurrences of these behaviours identified within the coding process.

User behaviour: Scanning

Scanning refers to the behaviour where the user surveys the content on the page before taking any action (Hawk and Wang in Hsieh-Yee, 2001; Kim et. al., 2012; Xie, 2010). 65.4% of all user behaviour was shown to consist of scanning, making it the most prevalent interaction behaviour used in the think-aloud observations.

User behaviour: Search and Browse

Carmel et al. (1992) argue that it is difficult to distinguish between search and browse as the two are generally considered overlapping concepts. In this context, search refers to the action where the user has already identified a source of interest and is specifically looking for that source when using the interface. For example, P1 was already familiar with the AP+P collection, and so when loading the *Works and Networks* interface they immediately started looking for a particular artist, rather than approaching it with no target in mind. Browse is identified as the activity of semi-directed search in areas of potential interest (Choo, 2000). Search and Browse were the second most prevalent form of user interaction behaviour, with 23.1% of the total.

User behaviour: Methodical

Methodical user behaviour occurred when the participant followed a systematic pattern as they absorbed information within the interface, for example, when using *Decade Summary*, P1 would hover over a thumbnail, click to see the larger artwork and then repeat the process, systematically moving through the data. Of the segmented user

interaction behaviours, 7.7% of these interactions were methodical, making it the third most prevalent form of user interaction behaviour. It was most apparent when participants were using the *Works and Networks* interface.

User behaviour: Chaining

Chaining is the name given to the process a participant undertakes when they follow a logical progression of links through a site (Ellis in Meho & Tibbo, 2003; Choo, 2000). It can be backwards or forwards, as Choo (2000) explains: "backward chaining takes place when pointers or references from an initial source are followed. [...] Forward chaining identifies and follows up on other sources that refer to the initial source." Backward chaining is much more common than forward chaining and is "a well established routine of information seeking among scientists and researchers" (Choo, 2000). Chaining was the least popular of the user interaction behaviours, with only 3.8% of coded behaviours.

User experience (umbrella term)

User experience is the umbrella term for categorising the types of experience the participant had whilst using the interfaces. There were four sub terms: enjoyment and surprise or delight which were considered to be positive, and annoyance and confusion, which were negative. Of the 118 coded segments, 52 (44.1%) were positive and 66 (55.9%) negative.

User experience: Enjoyment

Enjoyment is a superordinate term, defined as "the pleasure you get from something" (OUP, 2013). Seligman and Csikszentmihalyi (2014) describe how enjoyment refers to the good feelings people experience when they do something that stretches them. For example, a user can enjoy the experience of using an interface because it will create positive feelings. This was demonstrated by P3, who when using *All Artists* remarked:

> "Oh and you can filter by other stuff too? That's great. Yeh, this is really sweet." [P3:33]

24.6% (29) of coded behaviours were from users expressing enjoyment throughout the think-aloud observations.

User experience: Surprise or Delight

In section 3.5 I foreshadowed how feelings of surprise or delight are central to the notion of serendipitous discovery. They are two terms that are closely related, surprise is defined as "an unexpected or astonishing event" (OUP, 2010) and delight is a feeling of great pleasure (OUP, 2010). It can be hard to distinguish between them, and so I encoded them together. In the following example, feelings of surprise or delight arose from a user discovering an artwork they really liked (in this case, when seeing an artwork by Arthur Wicks in *Works and Networks)*, P1 exclaimed:

> "wow, these are really lovely aren't they?" [P1:95]

The feelings of surprise or delight were encountered in 23 (19.5%) of the total coded behaviours

User experience: Confusion

Confusion refers to any instance where the user was uncertain about an aspect of the interface or unclear about how it functioned. I saw a clear difference between Confusion and Annoyance, as a user could be confused by the way something worked in the interface, but that may not have necessarily annoyed them. Therefore I encoded them separately.

Of the total coded user experience segments, 26.3% were from the participants expressing confusion about aspects of the interfaces or actual site content. It accounted for 47% of all the negative experiences.

User experience: Annoyance

This is a general phrase used to categorise any annoyance that occurred whilst using the interfaces. Most annoyances were technical, for example, where the page would not load or the interface did not work as expected. It was the most significant user experience, comprising 29.7% of the total and 53% of all the negative feelings.

General feedback

Participants also provided five comments about the *Explore* interfaces in general that will be discussed later in this chapter.

Web survey

The Web survey was open for six months from the end of December 2013 until June 2014, during which 24 participants provided feedback. The vast majority of respondents were between the ages of 25 and 44 (15); 6 participants selected the 45 – 64 age range and only 3 indicated they were over 65 years of age. 54% of participants were male and 46% female. All the participants rated their average internet experience at 3 stars or higher (the average was 4.38), and 42% of participants stated they were expert internet users.

There were 59 unique visits to the survey using a PC or laptop but only 23 responses, indicating a completion rate of 39%. 5 people using a tablet visited the survey and only 1 of these visitors actually responded. The average time for completion was 4 minutes and 10 seconds on a PC and 6 minutes and 26 seconds on a tablet.

There were 75 segmented comments with 3 main codes and 13 subcodes. The main codes were already determined by the structure of the Web survey. They consisted of: positive comments (54.7%), negative comments (37.3%) and suggestions (8%). The 13 subcodes will be reviewed in the detailed discussions later in this chapter.

Data logging

Google Analytics has been installed on the AP+P website since it was launched in April 2013. For this discussion the GA data has been narrowed to align with the same dates of the Web survey. During the six month period 88,000 unique pageviews resulted from 34,279 sessions by 28,002 users to the entire AP+P website. Of these 3552 were unique pageviews on the *Explore* interfaces (*Decade Summary, All Artists, Works and Networks*) from 2762 users in 380 sessions.

Analysis of the GA data shows that *Works and Networks* was by far the most popular interface, with 2305 unique pageviews, *All Artists* had 799 and *Decade Summary* 448. The average time spent on *Works and Networks* was 54 seconds, on *All Artists* it

was 2 minutes and 30 seconds and 2 minutes and 29 seconds on *Decade Summary*. The average time spent on the whole AP+P website (excluding the *Explore* section) was 43 seconds.

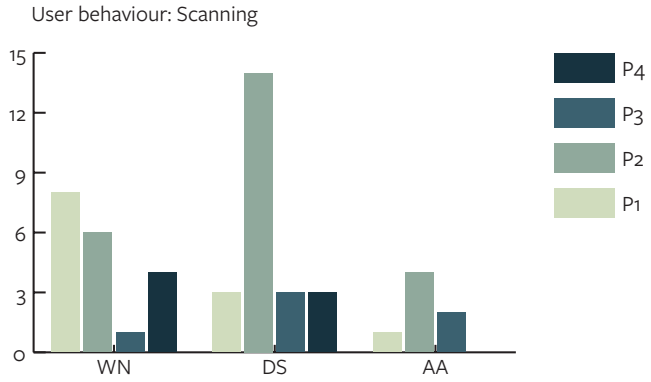
Detailed discussion of results

The discussion that follows is grouped by the research method, starting with the results from the think-aloud observations, followed by the Web survey, and finally the data logging.

Think-aloud observations

The goal of the think-aloud observations was to gain insight into how participants used and engaged with the interfaces. The findings are presented below.

User behaviour: Scanning



Graph 1
User behaviour: Scanning use per participant

WN	*Works and Networks*
DS	*Decade Summary*
AA	*All Artists*

Scanning was most prevalent when using the *Decade Summary* interface, followed by *Works and Networks* and *All Artists*. The participant who used scanning the most was P2; followed by P1; P4 and finally, P3 had the lowest rate.

Scanning was most apparent in the observation of P2 and P1, who took the time to read everything on the page before making a choice. It became clear that it is through the scanning process that the user learns how the interface works, for example, P2 remarks:

> "I like how it's very obvious after a moment that it [*All Artists*] is colour coded, because I think that it's very easy to see that [when] just looking at a field that green is going to stand for a gallery or a group." [P2:42]

> "this is quite cool [referring to the thumbnail display in *Decade Summary*], I don't know how easy it is to navigate but [if you are] just having a browse then it's really great... as someone who doesn't know anything about the content I could potentially spend a lot of time [looking at this]" [P4:65]

Scanning was prevalent when looking at large overviews, for example P2 selected the 'printer' filter in *All Artists* and then scanned the data [P2:117-121] before remarking:

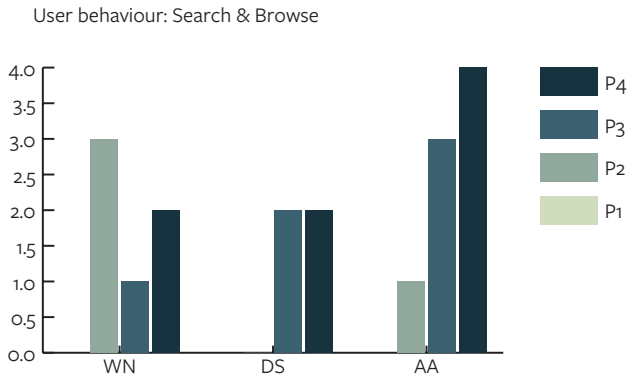
> "ohhh, so I start looking, [it] starts to get interesting" [P2:122]

Additionally, P3 spent time scanning the overview of *Works and Networks* [P3:101-114], before interacting with it. In these cases there was a lot of different information for the participants to observe. Once the participants had digested the information displayed, they tended to follow a more methodical process; a further analysis of this behaviour follows.

User behaviour: Search & Browse

Graph 2
User behaviour: Search & Browse use per participant

WN *Works and Networks*
DS *Decade Summary*
AA *All Artists*

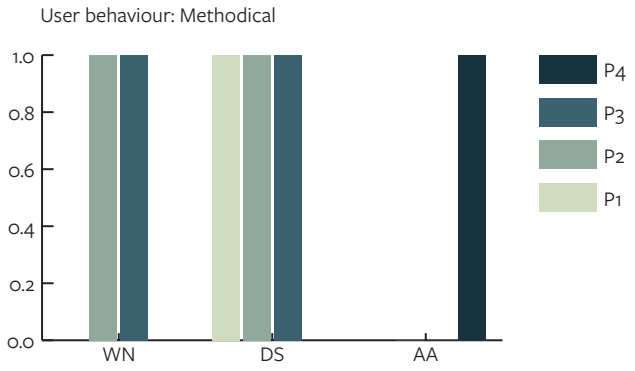


This user behaviour occurred most prominently in the *All Artists* interface, followed by *Works and Networks* and *Decade Summary*. Use of the behaviour by the participants varied. P1 was not identified

using it, whereas P4 used it eight times; P3 six times; and P2 four times. For instance, when using the *Works and Networks* interface, P2 tended to search for certain artworks or artists within the information shown, remarking:

> "there's a limit to how much I can actually use [the interface] because... there's no way I can alter the middle pane" ... "What I like to do when I get to someone like Bea Maddock is I know it's a big field and I know there's this great artwork but I want to find it and I can't narrow down the results from here" [P2:229]

User behaviour: Methodical



Graph 3
User behaviour: Methodical use per participant

WN *Works and Networks*
DS *Decade Summary*
AA *All Artists*

Methodical user behaviour was noticeable when participants were using *Decade Summary*, it was followed by *Works and Networks* and *All Artists*. The behaviour was spread consistently across the participants, both P2 and P3 engaged in methodical behaviour twice, while P1 and P4 did so only once.

This user behaviour generally occurred after the participant had digested the information displayed (for example, through the scanning behaviour above), they then tended to follow a methodical pattern of interaction. This is demonstrated by P2, who repeatedly undertook this process in *Works and Networks*:

> Hover over artist name > Click artwork thumbnail > view larger artwork > click on an artwork to view the full catalogue reference (in new tab) > click artwork thumb to the full size artwork [return to *Works and Networks* and repeat] [P2:144]

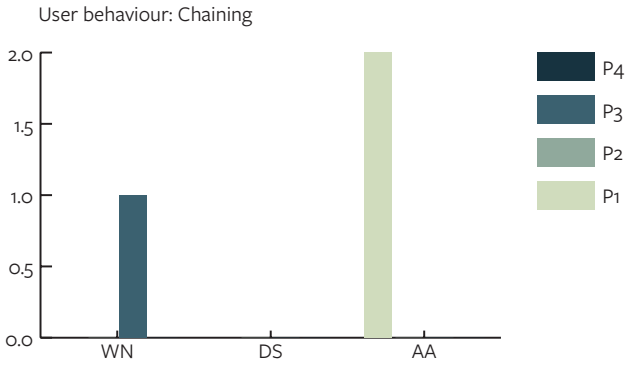
For P1, this was a common pattern when using the *Decade Summary*:

> Hover over a cropped artwork image to see the slightly larger uncropped artwork in the tooltip (however, not actually clicking on the artwork to load the even larger artwork) [and repeating] [P1:174]

User behaviour: Chaining

Graph 4
User behaviour: Chaining use per participant

WN *Works and Networks*
DS *Decade Summary*
AA *All Artists*



Chaining was the least popular user behaviour, with only three coded instances. Two of these occurred within *All Artists*, both from P1. The other instance was attributed to P3, whilst using *Works and Networks*.

It must be noted that both the methodical and chaining user behaviours often occurred at the same time. For example, P1 engaged in a backward chaining process that was initiated through a methodical interaction. They would start at a singular point of reference (an artwork) and then systematically follow the various links from the *Explore* section, to the catalogue reference on the main AP+P site, and then repeat the process again. The chaining pattern was:

> Click on a artwork thumbnail > click the artwork image to load the full artwork information page > click the artwork to view the full size artwork image > click the artist's name > view all their artworks [P1:208-213]

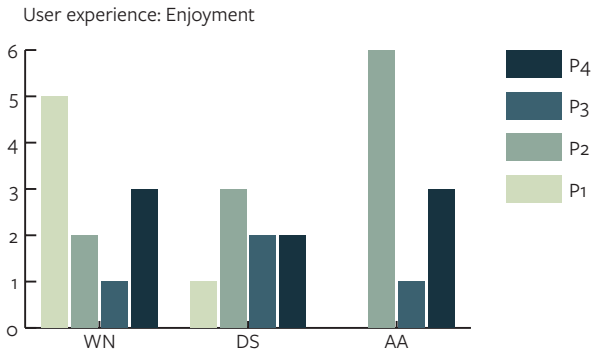
P3 would follow a very similar process, for example, from *Decade Summary*:

> Click a artwork thumbnail > click the larger artwork image to load the full information page > click related records > click on an artist name > click other related record information [P3:70-90]

It's a process they enjoyed, explaining:

> "This is good, I mean I like that it's got a lot of stuff, like all the metadata that you would want for expanding your search and finding related things." [P3:90]

User experience: Enjoyment



Graph 5
User experience: Enjoyable experiences per participant

WN *Works and Networks*
DS *Decade Summary*
AA *All Artists*

Of the 29 enjoyable experiences, 37.9% were encountered in *Works and Networks*; 34.5% occurred in *All Artists* with the remaining 27.6% from *Decade Summary*. P2 experienced the highest number of enjoyable moments, the majority (6) of which were whilst using the *All Artists* interface. P4 had the second highest number (8) of enjoyable experiences 3 of which occurred in *Works and Networks*, 3 in *All Artists* and 2 in *Decade Summary*. This was followed by P1, with a total of 6 enjoyable experiences. 5 of these occurred in *Works and Networks*, 1 in *Decade Summary* and none in *All Artists*. Finally, P3 had a total of 4 enjoyable experiences: 2 in *Decade Summary* and 1 in each of the other interfaces.

Participants' feelings of enjoyment from aspects of the interface are evident in the following comments:

> [in *All Artists*] "the width of these... this is the amount of artworks that you guys have of theirs... ok that's cool"[P4:154]

[hovering over artist list in *Works and Networks*] "oh, but if you look at that, it's highlighting that it was also printed by Martin King. That's cool!"[P2:139]

[in *All Artists*, click an artist box] "oh what's that? what happened there? [hovered over artist link and the *Works and Networks* tooltip shows]. Click to open, that's a great little tool because that's exactly what I wanted to do!"[P2:128]

[In *All Artists*] "it's quite quick considering you can just pull up a bunch of thumbnails out of 4000 boxes]"[P4:7]

There was also the sense of enjoyment expressed through comments about the actual design of the interfaces:

[W+N] "It looks great" [P2:223]

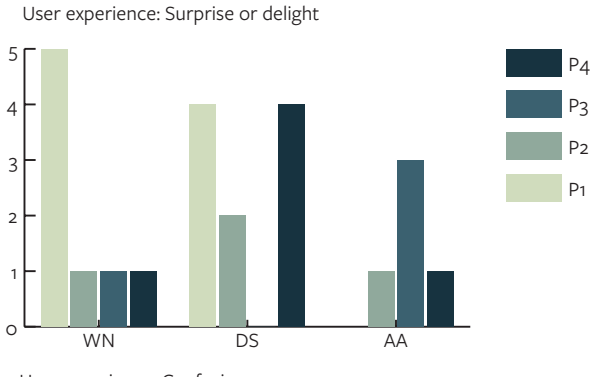
"It's good though... I mean there is so much complex information here and you have a lot of different ways to get into it which is good." [P3:64]

[*All artists*] "was mind blowing. It's crazy!" [P1:204]

User experience: Surprise or Delight

Graph 6
User experience: Expressions of surprise or delight per participant

WN *Works and Networks*
DS *Decade Summary*
AA *All Artists*



Expressions of surprise or delight were conveyed the most whilst participants were using *Decade Summary*, with 43.5% (10) of the coded instances. 34.8% (8) occurred in *Works and Networks* and 21.7% (5) in *All Artists*. P1 had the most surprising or delightful experiences with a total of 9 occurrences, 5 in *Works and Networks*, 4 in *Decade Summary* and none in *All Artists*. P4 had 6 total expressions of surprise or delight, 4 of which occurred whilst using *Decade Summary*. P2 and P3 both had 4 surprising or delightful experiences, for P3 the majority of these occurred whilst using *All Artists*.

Feelings of surprise and delight were expressed when the interfaces caused the participant to gain an unexpected insight:

> "I don't think you would ever come [across these] or see these if you searched in a normal way" [P1:113]
>
> "I was looking for something and I ended up somewhere completely different and it was actually something really good... that's cool isn't it!" [P1:116]
>
> "so it is kinda somewhat serendipitous" [P1:185]
>
> "who's that? I love this [referring to thumbnails]. I love scrolling over this and finding who it is. That often distracts me." [P2:253]
>
> "As someone who doesn't know anything about the content, I could potentially spend a lot of time [looking at this]" [P4:67]

P4 was surprised by how the interface changed when clicking on the print type segments in the *Decade Summary*:

> "Oh and it'll [the thumbnail box] change for each year right? yeh cool, yeh actually I really like that" [P4:72]

Regarding, *All Artists*, P2 and P3 were delighted by the functionality of the interface:

> "Oh I can see lots of people [at once]? And I can scroll through? oh that's nice." [P2:74]
>
> [when using filters in *All Artists*] "Oh I like that, oh and you can filter by other stuff too? that's great! Yeh, this is really sweet" [P3:24]

Interestingly, participants also expressed surprise or delight when the interface did not function as a more traditional search based interface would. For example,

> "[It's] not so good for research because you can't narrow it down yet, but for exploration it is nice" [P2:187]
>
> "But I don't want everyone, its not useful for me to have every Australian artist because I don't think you can look at something that big" [P2:313]
>
> "I'm always looking for things that will help me do research and it'd be great if I was able to get this window [the thumbnails in Decade Summary] and narrow it" [P2:305]

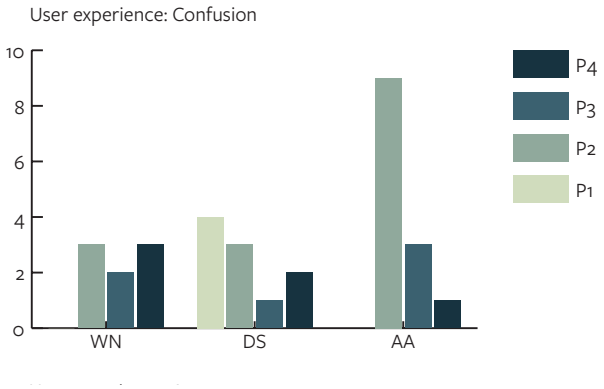
"I was thinking of the name of an artist, but that's not the function is it? You can actually go in there and have a whole range of ideas" [P1:84]

[When using Works and Networks] "It could be the starting point but you might be looking for something in particular" [P1:107]

User experience: Confusion

Graph 7
User behaviour: Moments of confusion per participant

WN *Works and Networks*
DS *Decade Summary*
AA *All Artists*



53% of the negative experiences came from the participant expressing a form of confusion. P2 had the highest rate of confusion, followed by P3 and P4, and finally P1. *All Artists* caused the most confusion (13) followed by *Decade Summary* with 10 and *Works and Networks* 8 confusing events. These negative expressions were manifested in different ways.

P1 did not understand the distinction between the interface (i.e. how it looks and functions) and the actual underlying data that it is built upon. When using *Decade Summary* they were worried by the lack of information from 2010 onwards:

"I thought this was really interesting: why does it stop in 2010? Is it because the money ran out?... it looks shocking, it's like why? what happened?" [P1:161]

The reason it stops at 2010 is because this was when the last batch of data was added to the AP+P site. Artworks are still being added to the collection, but the data is yet to be available.

P2 was particularly confused by parts of the *All Artists* interface, specifically the choice of colour to symbolise gender:

> "why is man blue? blue stands for man? red stands for female? that's interesting" [P2:46]

And P2 had difficulties understanding how the filters work:

> [click filter] "wait, what just happened there?! [click a new filter] What? What's happening?!... It would be good if the whole thing rearranged " [P2:119]

The filters also confused P3, who remarked:

> "Oh that's a bit confusing... I went here [checked filter] and there is no back button, but if I click the actual back button then I go back to the previous page [rather than removing the filter]... maybe there could be a clear all filters button?" [P3:6]

Two participants were unsure about what the width of the artist's name box was relative to:

> [the width of the box] "is this the length of their life or the length of their art?" [P3:19]
>
> [in *All Artists*] "I find it kind of interesting that it's not pointed out that this is only the NGA's collection, so it [the data] isn't representative of exactly how big their output is, because Frank Brangwyn is a huge printmaker and his box is tiny" [P2:33]

The width of the box is relative to their total artwork count across all roles. In this particular case, Brangwyn[6] is listed as the printer for only one artwork in the AP+P collection; he may have produced more artworks but they are not represented in this dataset.

Works and Networks was considered the most confusing interface, with both P3 and P2 being unsure about the terminology used within it. P3 took issue with the word 'Networks':

> "What do you mean by networks? Is that how people are related to each other? Because I don't see any actual networks here. I would think it'd be more like a word cloud" [P3:98]

6 http://printsandprintmaking.gov.au/explore/works-and-networks/?artistid=4393

P2 was unhappy about how the artworks without any date information were abbreviated as 'n.d', they explained:

> "to have a thing that says no date is... a bit misleading... It seems to me it's like putting up a wall to people who are trying to find out information. If it were unknown date, that would be a lot better" [P2:2]

The printmaking terminology used in the data and the interfaces is quite specific and so bound to cause confusion to those unfamiliar with the field. This is shown by P4, who asks:

> "what are binders?" [P4:47]

P1, who has a background in printmaking, seemed to pre-empt the question and wondered:

> "does it give you an idea about what the terminology is? like say if you were a complete novice and had no idea what planographic print was, or stencil, can it take you into a description of terms?" [P1:188]

Finally, there was general confusion about the navigation:

> "If I see that Fred Williams has 13 collab [collaborations] with George Baldessin then I'm going to press that, but if I do it sends me to that artist's page rather than highlighting any artworks" [P2:212]

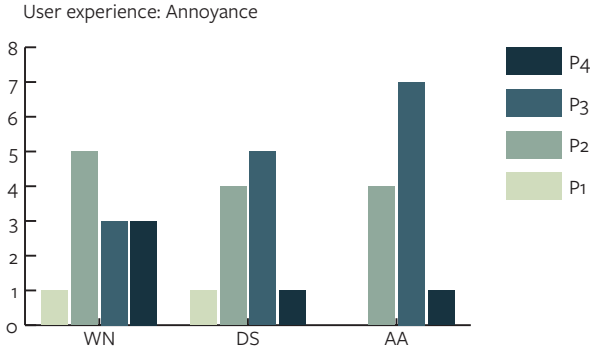
and how the data display was ordered:

> "Is this all chronological? [referring to the artwork thumbnails] No maybe not, or maybe it's reversed? It would be good to be able to sort by time" [P4:84]

User experience: Annoyance

Graph 8
User behaviour: Moments of annoyance per participant

WN	*Works and Networks*
DS	*Decade Summary*
AA	*All Artists*



User interface and interaction issues were the greatest cause of annoyance in the think-aloud observations. Both *Works and Networks* and *All Artists* had the same number of occurrences, whilst *Decade Summary* had one less. Of all the participants, P3 had the most moments of annoyance, with 15 in total, 7 of which came from *All Artists*. P2 had the second highest rate of identified annoyances, with 13 different events; followed by P4 with 5; and P1 with 2 annoyances one from *Decade Summary* and another from *Works and Networks*.

The main user interface problem was caused by the three-column layout, as P2 explains:

> "When you scroll over the artist's name, it is handy because it highlights the one's [artworks] that you are looking for. But I've found that you have to keep hovering back over their name because you can't scroll while looking." [P2:210]

P4 encountered another problem where they were unable to scroll down the page:

> "I can't get down there for some reason" [P4:51]

A major annoyance came from leaving the current view, whether intentional or not. For example, when P3 was using *All Artists* and clicked on an artwork, the detailed artwork view opens within the current browser window (rather than a new tab) which destroys the current user view. If they try to return using the back button the interface will load again at its default starting point. P3 remarked:

> "Is there anyway to save the viewing state? It's a hassle to lose where you are at" [P3:56]

Additionally, P2 did not appreciate how the artwork thumbnails were cropped in the *Decade Summary*:

> "If I was looking for pleasure, these [the thumbnails] are so hard because you can't see them" [P2:287]

General Feedback about Explore interfaces

Five general comments related to the *Explore* interfaces. *Decade Summary* received the most comments, all of which were positive:

> "Oh I Love this one! When I first found it I just played for ages, hovering over the different elements... I thought it was really interesting" [P1:154]
>
> "So the Decade Summary is great." [P2:267]
>
> "[It] is a really cool way to visualise lots of information which is obviously a challenge" [P3:203]

On *All Artists*:

> "That [*All Artists*] was mind blowing!" [P1:204]

Web survey

This analysis of results from the Web survey is split into general discussion about positive and negative aspects of the *Explore* interfaces (as a whole); and then targeted discussion about each interface.

The full response of each participant was also coded as being positive, negative or neutral (Graph 9). Some participants were very negative, for example P190, left two comments, the first in the 'enjoy' field:

> "they did not load – I tried several times. So unfortunately I could not enjoy anything." [P190]

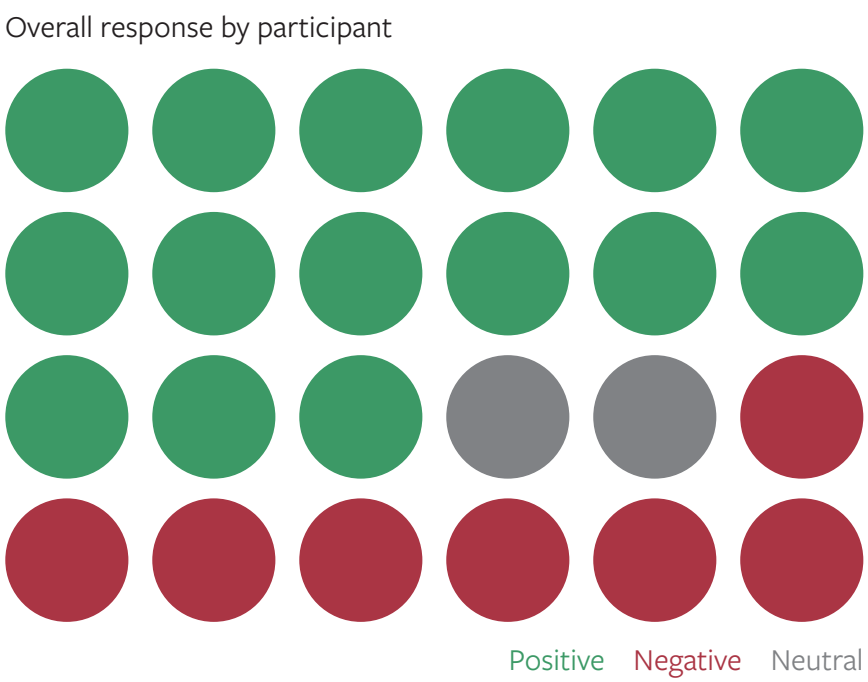
And the second in the 'did not enjoy' field:

> "poor website performance" [P190]

As these comments both regard technical matters, I assume the participant was using Internet Explorer to access the interfaces. There were no recorded server outages during the period of the study and they load on all other browsers and devices.

Whereas the majority were very positive, with P175 commenting:

> "This is hands down the best interface of its kind I've seen" [P175]



Graph 9
Overall responses by participant

General comments

The general comments were initially encoded as being positive, negative or offering suggestions. They were then further segmented into more detailed subcodes (Graph 10).

Positive comments

Four main sub-codes emerged when analysing the positive comments given in the Web survey: comments on the whole project; on exploration; and on the design and speed of the site (Graph 11).

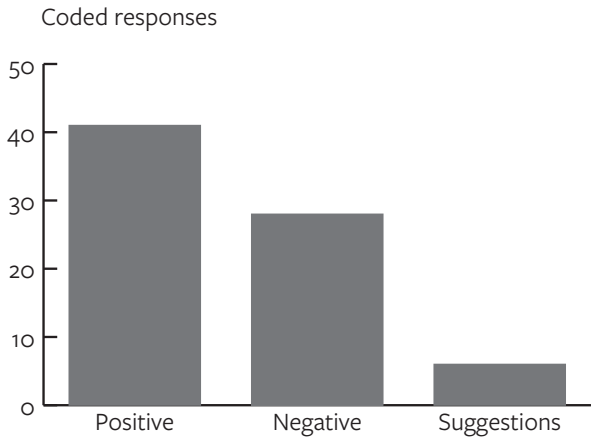
Comments about the whole project consisted of several enthusiastic responses like:

> "I thought it was great" [P167];
>
> "it is a great resource" [P154];
>
> "overall, love it! and learned from it!" [P215]

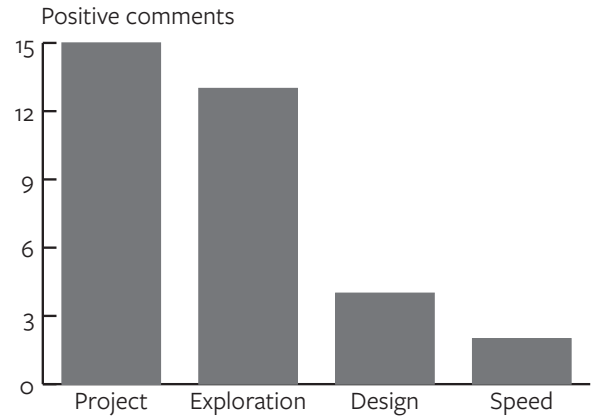
There was also encouragement about the prospects of further development:

> "great project, can't wait to see more generous interfaces and new ways of interacting with collections" [P167]

Graph 10
Coded responses



Graph 11
Positive comments: sub-codes



> "A terrific and highly accessible resource, would love to see it grow and expand in the future" [P213]
>
> "keep up the great work!" [P161]

One comment referred to the importance of linking between the interfaces, and back to the full reference in the AP+P website:

> "great cross-referencing, learned a lot about Oz art along the way too" [P192]

Comments regarding exploration included specific responses to the process of exploration within the interfaces:

> "you can wander through without knowing anything about prints and printmaking" [P180]
>
> "I can just browse the collections with no specific object in mind. It allows me to engage with objects that I would not otherwise even look at." [P186]
>
> "visually engaging, informative, opportunity for exploring rather than static searching" [P217]

> "love the different ways of navigating and different layers" [P215]

Feedback on the design showed that participants enjoyed the minimalistic style of the interfaces:

> "[I liked] its design, style, modernity" [P169]
>
> "I also like that you avoided flashy, but non useful, animations" [P175]
>
> "overall feedback: is that the strong visual component artworks very well, and that the interface is simple, clean, powerful and intuitive" [P175]

This participant appreciated how the three interfaces each provide a different view of the collection:

> "visual appearance: the fact that there are three ways of looking at the same collection emphasises different aspects and allows one to discover different things" [P148]

Finally, general feedback about the **speed of the website** was given:

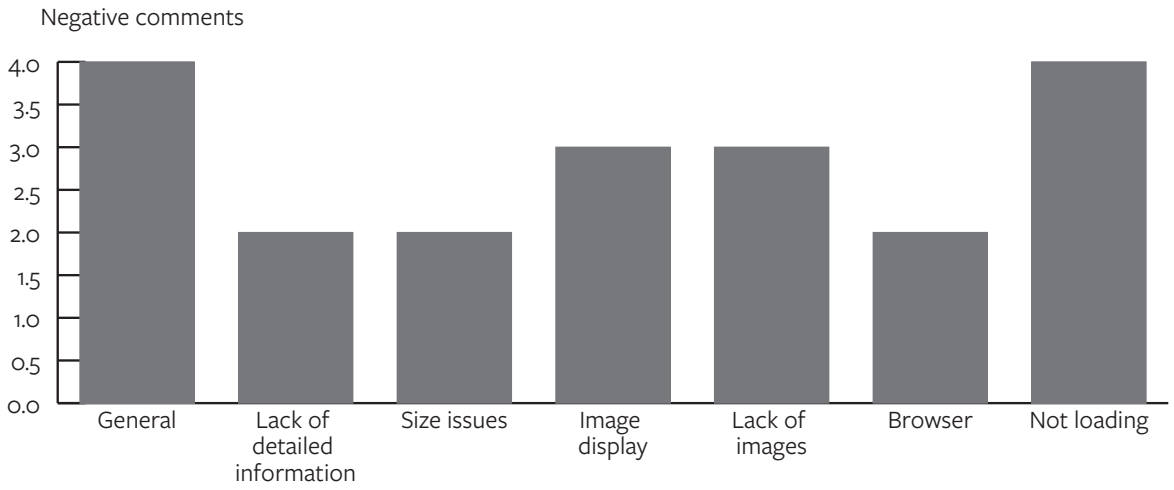
> "I truly appreciated that however you implemented this, it is not real time intensive, and does not bog down my browser" [P175]
>
> "it is fast" [P154]

Negative comments

Seven subcodes were identified within the general negative comments: feedback on the whole project; lack of detailed information; image display; lack of images; display size; type of browser and loading issues (Graph 12).

Feedback on the whole project included remarks about how the interface is accessed:

> "the interface could just open once when you open an artist's artwork" [P154]



Graph 12
Negative comments: sub-codes

Frustrations about the lack of search:

> "the difficulty of searching for relevant information" [P196]
>
> "if I was after a specific artwork it would be frustrating to navigate through to it" [P167]

Two participants were annoyed by the **lack of detailed information** displayed within the interfaces:

> "maybe being able to see more of the metadata straight away (when you select a print) could be useful?" [P167]
>
> "I would like more content information. I always find basic catalogue information a little dry" [P186]

Many comments concerned the **images**, from how they were displayed:

> "the size of the thumbnails is a bit weeny" [P162]
>
> "It was a little frustrating to quickly bring up larger images" [P213]
>
> "frustrated by size of thumbnails – they were too small" [P215]

To the apparent **lack of images**:

> "all the x's indicating no images – although I am impressed by the number of images that do show up" [P175]
>
> "not enough images" [P192]
>
> "sometimes not all the pictures were visible" [P223]

Additionally, there was feedback about what **browser** could be used to view the interfaces, and how that would change the experience:

> "I'm not sure if there is full functionality on an iPad, but it's the kind of site that would be best to explore on the couch rather than at a desk" [P180]

> "It would be nice if it also worked through Internet Explorer" [P223]

Finally, a number of participants expressed frustration about how the interfaces loaded, particularly regarding the page speed:

> "the slowness of loading, the constant unresponsive scripts" [P196]

> "sometimes the pages took a while to load and/or were laggy" [P219]

Two participants experienced the ultimate **failure**:

> "It promises much, but does not yet deliver. No interfaces able to be opened" [P184]

> "they did not load, so unfortunately I could not enjoy anything" [P190]

Comments about individual interfaces

All Artists

No specific positive comments were received about *All Artists*. However, there were three negative ones:

> "[I don't like] the fact that most links open new windows. Having an area dedicated to detail view as in the other two interfaces would streamline the experience" [P158]

Interestingly, two comments were made about issues that were also identified in the think-aloud observations. The first concerns the colour choice for different genders:

> "red for female is most visible, blue for male is much harder to see – maybe make all colours equally saturated" [P215]

Secondly, there was confusion about how the filters worked:

> "I found left hand nav [navigation] a little confusing. I figured out that you click the box to get that category. But you can also click the name of the category which yields different results" [P215]

The filter name is attached to the checkbox, which means there is no way for different results to be returned. It is unclear what the issue this participant is referring to.

Works and Networks

Feedback about *Works and Networks* shows that participants appreciated the focus+context display, with the three column layout allowing three levels of data to be shown:

> "most links open in the current window, which offers a seamless exploration experience" [P158]

> "the way each entry offers a large list of related links, which helps explore the collections by association of ideas" [P158]

> "love the left nav buttons that reverse colour" [P215]

The layout did cause confusion for one participant who remarked:

> "[It] has not become intuitive for me. I suppose it's the relationships between agents, but it's not obvious to me what the third section does" [P180]

There was also uncertainty about how the *Works and Networks* interface loaded the starting workshop, with a participant describing this scenario:

> "When I clicked on [Works and] Networks just now, the Cicada Press page came up. I may have gotten the interface into a weird state, or maybe it is user error. I did get to the page with all the Networks listed previously. I clicked the back arrow in my browser, and then clicked [Works and] Networks again, and this time got Port Jackson Press." [P175]

This indicates that they did not understand that the starting workshop is randomised, which is logical as I never actually provide any instructions or explanations to clarify what is occurring. As another participant explained:

> "I didn't quite understand the basis on which the *Works and Networks* portal randomised the results – if it did do this?" [P219]

Decade Summary

Decade Summary appeared to be the favorite amongst the Web survey participants for a variety of reasons. Positive comments about the navigational histogram:

> "love the left-hand nav visuals" [P215]
>
> "the decade barchart felt good" [P162]

And its overall use:

> "*Decade Summary* by media is very useful" [P175]
>
> "I especially like the *Decade Summary*" [P219]

There were only two negative comments about *Decade Summary*, both of which regarded the histogram:

> "[It is] a bit busy, [and] slightly overwhelming as you scroll your mouse across the timeline" [P217]
>
> "[The] timeline breakdown of print types is hard to navigate, especially the small slices." [P161]

Suggestions

Finally, six suggestions were provided by four participants. They ranged from improvements to the current interface or new features that could be added to them. For example:

> "Additional view [in *Decade Summary*] of single decade timeline above 'results" pane may work better" [P161]
>
> [in *Decade Summary*] "wished I could select a category to view e.g. reliefs so as to see the reliefs across decades" [P215]
>
> "some linking between the 3 interfaces could be implemented e.g. when looking at a single artwork in *Works and Networks*, ability to switch to *Decade Summary* by clicking on the type-year information under the title" [P158]

Ideas about the development of further interfaces were as follows:

> "I would love to see further filtering options based on hues/colours and other interesting metadata that could be extracted" [P161]
>
> "I'd love to find a way to see the *whole* collection. Somehow." [P162]

And following Shneiderman's Visual Information Seeking mantra more closely, one participant stressed:

> "making sure the emphasis is on the imagery first, then finer detail when required would be good" [P213]

Data Logging

Works and Networks

Works and Networks received 2306 unique pageviews during the evaluation period. By comparing the unique pageviews with those for the other interfaces it is immediately clear that it was the most popular of the three.

Works and Networks functions differently to the other two interfaces: instead of being a Single Page Application, it loads new content in much the same way as a normal website. That means that if a user clicks on a related link then the id number for the artist is appended to the URL, for example, '/explore/works-

Artist name	URL	Unique pageviews	Avg. time on page	Work count
Australian Print Workshop	/works-and-networks/?artistid=14162	12	00:03:29	2715
George Baldessin	/works-and-networks/?artistid=6252	6	00:05:54	373
Greg Waddel	/works-and-networks/?artistid=18317	10	00:00:25	0
Rod Ewins	/works-and-networks/?artistid=5836	5	00:13:19	154
Barbara Hanrahan	/works-and-networks/?artistid=883	6	00:07:02	449

Table 5
Pageviews and average time in *Works and Networks*

and-networks/?artistid=14162'. It is important to be aware of this because it affects the way the analytics data is interpreted.

Of the 2306 unique pageviews to the *Works and Networks* interface, 540 were to the top level page (/explore/works-and-networks/) and 1766 to individual artist views within the interface. The total average time on page for both the top level and sub views was 56 seconds, however, this time period varies considerably when broken down into more specific views. See examples in Table 5.

Besides the extreme outliers, there appears to be some correlation between the amount of time spent on the page and the number of artworks displayed; this aligns with the results recorded in the think-aloud observations where scanning occurred within *Works and Networks* in 37% of coded occurrences.

An event tracker attached to the artwork thumbnail click shows that 1793 events were triggered by users clicking on an artwork thumbnail, in order to view a larger image. Of these events, 793 occurred on the top level page, indicating an average of 1.4 events per user. On the sub pages there was an average click rate of 0.56. Although it is not possible to break the data down by user, there appears to be evidence of users digging deep into specific artist's artworks, for instance, there are nine unique pageviews for Barbara Hanrahan with 66 events, indicating an average of 7.3 events per user; for Barbara Brush the number is significantly higher, with 2 unique pageviews and 28 events.

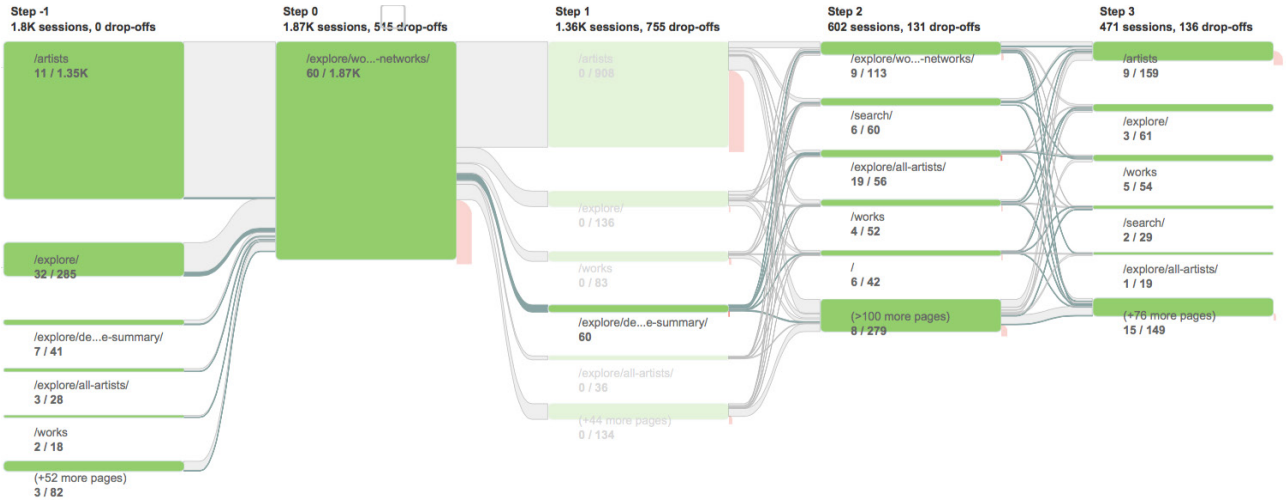


Figure 62
Behaviour flow for *Works and Networks*

It looks like the click pattern is random, but it is impossible to determine whether this is the case as users may be clicking on artworks they know or like.

Analysis of the behaviour flow data shows how users move between the different sections of the website, for example, from the interfaces to the catalogue references and back again. The data shows that the vast majority (1870) of users accessing the *Works and Networks* interface come from the artist details page on the main section of the AP+P website. On this page there is a link directly to the current artist within *Works and Networks*, the link is unique to the page and also the interface (the other two are not linked via the artist page – only in the footer).

Links are provided to access *Works and Networks* from within the other two interfaces too. This cross-linking is important as it allows the user to move between the interfaces. The behaviour flow establishes that 41 users came from *Decade Summary* and 28 from *All Artists*. Once they left *Works and Networks*, many users viewed the individual artist pages (908 users); some viewed detailed artwork information (83); and 60 visited the *Decade Summary* and 36 went to *All Artists*. Further investigation of the data shows that some users navigated through the various interfaces, moving to detailed information and back to the interfaces. This is demonstrated in Figure 62 which highlights behaviour flow between *Works and Networks* (step 0) and *Decade Summary* (step 1) and other pages, as step 2 indicates, some users return to

Works and Networks whilst other visit *All Artists* and so on.

The behaviour flow data clearly shows that users will navigate through different interfaces if given the chance. Some take as many as 11 different steps as they move through the website's different sections.

All Artists

All Artists was the second most popular interfaces during the evaluation period with 799 unique pageviews and an average page time of 2 minutes and 30 seconds. As previously mentioned, *All Artists* and *Decade Summary* are Single Page Applications and function differently to a normal webpage. To track how they were being used, custom event tracking code was added to various components of the website to allow for more fine-grained data collection. In *All Artists* these took the form of seven different events: toggle filter (sub split by role, count-band & gender); sort; view artworks; view more artworks; view previous artworks; close artworks by name and close artworks by 'x'. Table 6 lists the event information and counts for each.

A total of 7094 total events occurred, an average of 8.9 events per pageview. There were 715 events from the toggle filters. Table 7 contains a breakdown of the top five. The least popular was the 'no sex' toggle, with only 7 events (1% of the total).

Of the 51.8% or 3677 view artwork events, only 37 artists were clicked on more than 10 times. It is very revealing to look at the top ten artists and their placement in the interface (Figure 63), seven of the ten most popular artists are in the top two rows of the interface and numbers two and eight are the first two largest boxes on the screen.

Number six is George Baldessin and he is much further down the page.

Event name	How it is triggered	Count	%
View artworks	On artist name	3677	25%
More artworks	View more artworks	1011	14%
Close artworks (x)	On 'x' button	800	11%
Toggle filter	Filters	715	10%
Close artworks (box)	On artist name	664	9%
Previous artworks	View previous artworks	200	3%
Sort	Selecting drop down	27	1%

Table 6
All Artists event triggers

Event name	Note	Count	%
Role: artist		243	34%
Role: printers		119	17%
Sex: female		57	8%
Sex: male		40	6%
Countband-6	display 50-100 artworks	30	4%
Role: print workshop		30	4%

Table 7
Top five events in *All Artists*

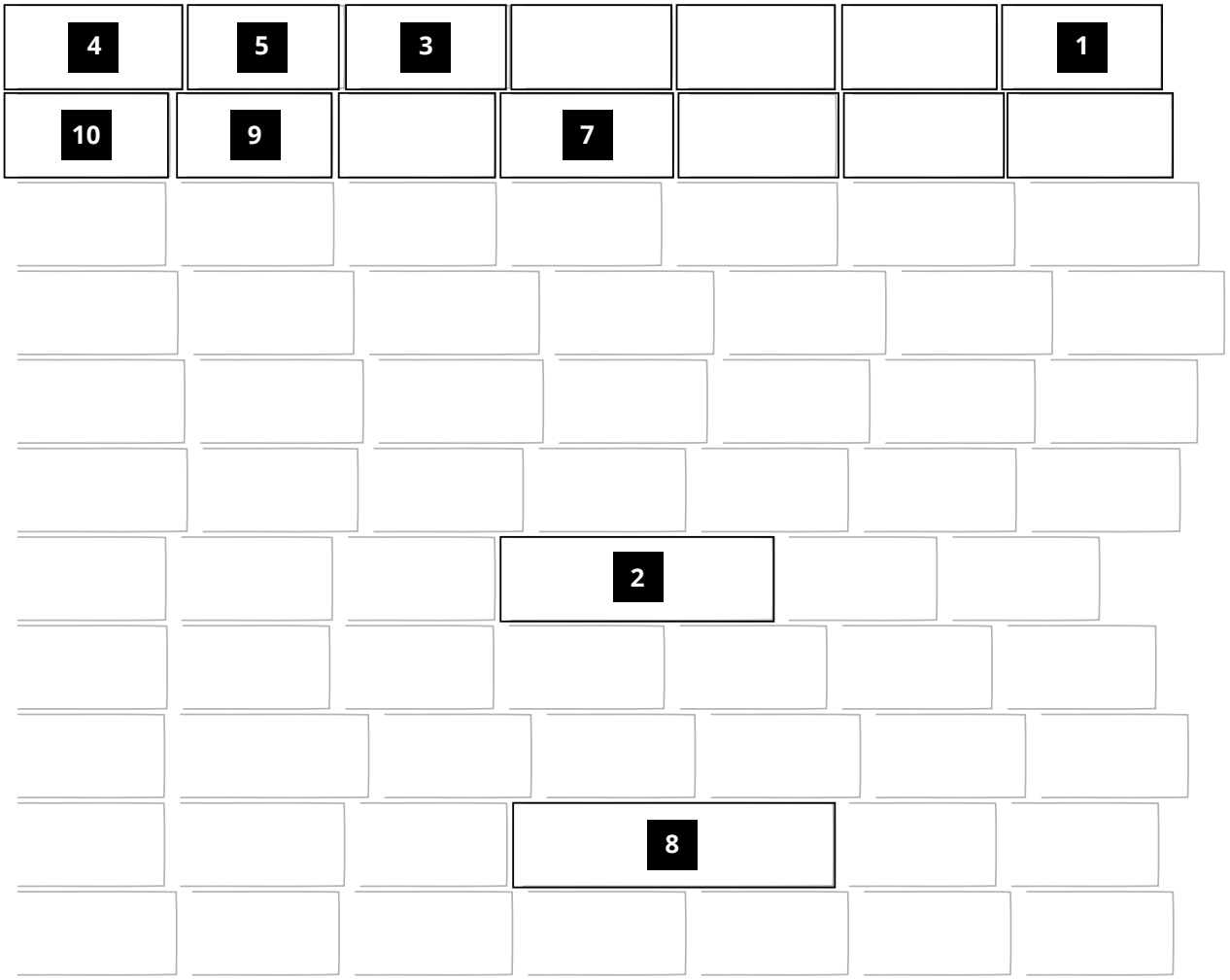


Figure 63
Artists with highest click rate

In regard to accessing the interface, a different pattern is evident to that of *Works and Networks*. In this case, the majority (384) of users accessed *All Artists* directly from the *Explore* page; an additional 36 users came from *Decade Summary* and 34 from *Works and Networks*. An additional 133 users came from a much wider range of pages—indicating they accessed the interface by clicking the direct link in the footer. The user behaviour flow data shows that users of *All Artists* tend to leave the interface by accessing further information on an individual artwork. In Figure 64 we see how the user behaviour flows through *All Artists* (step 0) to the detailed artwork information page at step 1 and then back into the other interfaces, the specific artist page or other parts of the site (step 2).

All Artists is the most browser intensive of the interfaces. It loads the data in large segments and displays a loading icon until it has loaded everything it needs. This process can be slow,

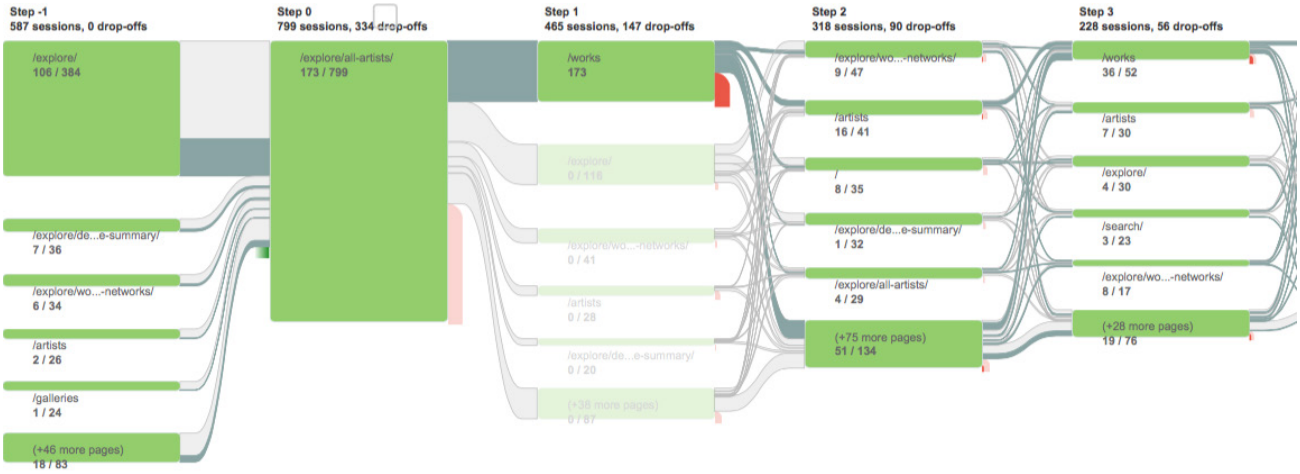


Figure 64
Behaviour flow for *All Artists*

as indicated in the page timing data, which shows an average of 21.82 seconds to fully load. The average page load time for the entire site (excluding the *Explore* section) is 2.39 seconds.

Decade Summary

Decade Summary was the third most popular interface during the evaluation period with 448 unique pageviews and an average time on page of 2 minutes and 29 seconds. As with *All Artists*, it is considered to be a Single Page Application, and so interactions within the interface were captured through the use of custom events. In *Decade Summary*, there were two main types of events: open big artwork and change segment. Together there were 3078 events, 65.53% of these were from clicking to open the big artwork and 34.47% from changing the selected segment.

Of the open big artwork events, a coherent pattern does not emerge. Instead, as expected, the artworks clicked on seem to be completely random. The top 3 most popular artworks (each with 6 views) are Planographic prints, however, whilst the first one appears on the 1960's segment (that is open when the interface is loaded), the others are from 2010.

It is also obvious that there is no consistent pattern in the 1061 'change segment' events that occurred. The most popular of these is intaglio prints from 2000 with 37 events recorded, whilst there were 10 events that were only triggered once. My

hypothesis was that these would be the smallest segments, however, the data shows that some are and some are not, and so it appears that users are exploring widely across the range of segments. The wide distribution of the 'change segment' events is clear indication of successful exploration occurring.

As has been established with *Works and Networks* and *All Artists*, users obviously move between the interfaces and the detailed information pages at their will. The number of users (62) visiting the individual artwork information page from *Decade Summary* is much lower than in *All Artists* (173) and *Works and Networks* (83). However, many users (163) do move to another interface, either by using the links in the header and footer or through the *Explore* page.

Summary

The evaluation study provided many valuable insights into how the *Explore* interfaces are used. It revealed their strengths and weaknesses, as well as providing detailed feedback on the various aspects that users did or did not enjoy. The discussion that follows identifies some of the key results of the mixed-method evaluation study, and explores some of the challenges faced in subsequent work.

One of the research questions that informed the evaluation study asked whether *the user was able to engage in open-ended exploration and discovery.*

Whilst addressing this question, it became evident that many participants enjoyed being able to wander through the collection in its various states. They liked the option of clicking on artworks of interest and viewing more detail within the same display. In the think-aloud observations, participants often found artworks that attracted them, or artworks they had not seen before, and followed the chain from the interface to the catalogue reference on the AP+P site. Results from the data logging showed the dramatic difference between the average amount of time spent on the three interfaces (1 minute and 58 seconds) compared

to 43 seconds on the non-explore section of the AP+P site. This clearly showed the strength of the interfaces in allowing free-form exploration and serendipitous discovery to occur.

It is clear that overall, participants regarded the interfaces as an effective collection access tool. They expressed moments of enjoyment, surprise and delight whilst interacting with them and commented on their successes. As noted in the literature review, it is these feelings of surprise and delight, which are integral to the notion of serendipitous discovery. It seems that users were keen to engage with new exploration tools, even if they were unsure about how they functioned. Participants appreciated the various styles of overview created by the three interfaces, but often wanted more options to refine these displays. The integration of various data visualisation techniques seemed to work well, with participants starting to understand the consistency of the visual language (and techniques) deployed across the three interfaces.

It is useful to reflect on the distinction between the interface and the content, as these elements were the source of some confusion and annoyance. As the developer, I understand the distinct separation between the source data and how I have chosen to display it. However, for the user, there is no distinction between these elements—the interface is the content. This conflates various issues, for example, there were many comments regarding the lack of images or the display of more detailed metadata. The user sees these issues as being a limitation of the interface whereas I know they are caused by external factors regarding the content. For instance, I show all digitised images from the NGA and the API calls are limited in how much data they output. As a result, the inseparability of the interface and the content emerges as an important consideration when evaluating the effectiveness of this style of work.

Nevertheless, many of the confusing or annoying factors drew my attention to areas for improvement and resolution in future work:

Be clear. Participants determined that it wasn't obvious what some aspects of the interfaces did or how they functioned. Abbreviation of the unknown date field into 'N.D' caused substantial confusion and instead of just showing an 'x' to indicate a non-imaged artwork, an explanation would be more useful. In *Works and Networks* participants struggled to understand that the starting point was randomised, an explicit statement about what was occurring would resolve some confusion. There is no doubt that incorporating basic instructions and explanations about the interface's functionality would improve the user's ability to understand how it works.

Allow refinement. An overview is great, however, participants immediately wanted to be able to refine the overview in different ways. For example, by changing the sort order or by filtering using extra metadata information such as location or colour (if available).

Consistency is important. Links in the interfaces all need to work in the same way—they should open in a new tab so a user does not lose their current view—and must be consistent across all interfaces.

Don't break basic Web conventions. Participants were annoyed when the interfaces didn't function as they expected them to. This issue was particularly apparent when there was no permanent URL for the current view, which in turn, meant that clicking the back button would reload the whole interface— rather than maintaining the previous view. This aligns with the 'history' stage of Shneiderman's Visual Information Seeking mantra, which states that a user should be able to retrace their view history.

Show larger images. Increases in screen size and internet speed have removed previous limitations regarding image sizes. Users have come to expect larger images and want to see them in these interfaces.

Some of these issues are more significant than others and I sought to address them in the development of *Subjects Explorer* and *Timeline*, my final two works.

Before turning to evaluate them, it should be noted that the initial *Explore* interfaces (*Works and Networks*, *Decade Summary* and *All Artists*) went through another rigorous form of testing through presentations I delivered at various conferences. Proposals submitted for these were part of a competitive peer-reviewed process. The international launch of the interfaces occurred in April 2013 at the *Museums and the Web* conference in Portland, Oregon where I gave a presentation entitled *Visual Exploration of Australian Prints and Printmaking*. They were first presented locally at the *Museums Australia* conference held in Canberra in May 2013.

6.4 Conclusion

In this chapter I have outlined the evaluation process which I undertook as part of my research project. The first stage of the process was crucial as the results directly informed the production of the final two interfaces. I demonstrated how a mixed-method approach allowed me to collect qualitative data, through the think-aloud observations and unstructured feedback in the Web survey, and quantitative data via the Web survey and data logging. Detailed analysis of the results showed that the three interfaces evaluated, *Works and Networks*, *Decade Summary* and *All Artists*, successfully engaged participants in the process of free-form exploration and serendipitous discovery. Participants appreciated the use of overviews and the data visualisation techniques and particularly, liked the detailed focus+context displays, as demonstrated in *Works and Networks*. The challenges raised by the evaluation study were also identified and discussed.

The following chapter will outline my contributions to new knowledge and reflect on the practice-based works through a discussion that addresses my research questions in depth. Finally, in chapter eight, I will summarise the project, its contributions to knowledge and discuss how the techniques I have developed are applicable to any digital collection.

7 Reflections

7.1 Introduction

This chapter provides reflections on the interfaces I produced and outlines the contributions to knowledge that have emerged through the practice-based research.

I begin by briefly describing my key contributions to new knowledge. I have created six unique interfaces which allow visual exploration of a large digital cultural heritage collection. This has been achieved through the development of new techniques for promoting exploration within a Web-based context. These techniques combine methods from data visualisation and modern Web design, and are applicable to any other collection of cultural data. I have evaluated the practice-based work and demonstrated how a mixed-method research study can provide valuable insights into this style of work.

In order to provide clarity within this chapter, I have grouped the discussion that follows into two key themes: technical development and exploration. Within these, I will discuss the key components of the project and the research questions that informed it.

In the first section of this chapter I return to the broad theme of technical development and reflect on associated issues. I will discuss the sketching in code methodology, before reflecting on the importance of structured data and how to access it effectively using APIs. I then consider the development of a consistent design language, the role of data visualisation and how it influenced the creation of new techniques for displaying cultural heritage data. To conclude the discussion regarding technical development, I reflect on how I utilised new Web-based technology to create the interfaces. The second part of the chapter discusses the theme of exploration and provides reflections on the concept of the overview, before identifying a distinction between the models of visualisation

and exploration and describing the need for techniques which can entice the user into the exploration process. I will then contribute further understandings about how to create generous interfaces by illustrating their application in my work, before assessing the role of evaluation. The final part of the chapter outlines my contributions to new knowledge in the cultural heritage field.

7.2 Technical development

My first research question asked "*How can digital interfaces that engage users in exploring cultural heritage collections be created*?" The primary concept here is concerned with the creation of the practice-based work, and the topics which follow are aligned with this focus. They emerged from a consideration of the individual research questions, the course of production and also incorporated reflections I made during the process of development.

Sketching in Code

In Chapter three, I outlined how I intended to use a sketching in code methodology. This would see the design and development of the interface occur at the same time, leading to a crossover of traditional development roles. It is an approach that required me to have a strong understanding of a wide range of different fields, from being able to conceptualise how an interface could look, to comprehending the multiple programming languages required to create it. It results in a development cycle which is very different from the normal process of a 'maker' where if you were making a chair, for example, you would know from the outset what the end goal is and would create many prototypes to find the most effective way to produce it. I could not know at the beginning whether the interfaces I wanted to create were possible because the project was totally dependent on the data.

Despite some of these drawbacks, I found the sketching in code methodology to be an extremely worthwhile one. The key to the success of the methodology is the use of real data from the very beginning of the project. This ensured that at every stage of

development the interface was a real thing—it either worked or it didn't. My development process consisted of a constant cycle of writing code, testing the interface in the browser, changing the code and testing again. With each iteration a problem would be solved and a new one would emerge. Whilst this can be frustrating, it is also a valuable process, as the development of a solution allowed me to immerse myself in the data, to understand what it contained, its structure and how to best approach it. It was through these constant iterations of the code that I was able develop the interface from a simple concept or idea into a fully fledged piece of work. Throughout my project, the form of the interface would always emerge through the process of production.

Structured Data

The quality of the highly structured data in the AP+P website was fundamental to this project. Nonetheless, translating the data from the database into a dynamic data-driven interface can be a difficult process and even the best data structures can have problems. In this project, I faced a number of issues, some were small and seemingly insignificant, for instance, extracting an ID number from a string; others were much more substantial, for example, dealing with unstructured data.

In Chapter four, I discussed a number of significant issues with the data that I had to overcome to produce *Subjects Explorer*. The final process of extracting the data was complicated and would be difficult to reproduce. Nevertheless, it is an important issue to reflect upon, as it shows a case where the structure of the data is not reflected in the data model. In other words, the subject's data exists in the database as a single string, however, within that string there are multiple subjects—it does have structure—but this isn't shown in the actual data model. The challenge here is to parse those strings and create a new dataset that contains the more structured data. The greatest problem with this technique is that the new subjects dataset is no longer 'live', it is based on a copy of the original dataset at a particular point in time. The data is now

separated from its canonical source, the subjects are not updated as new ones are added or amended and the artwork information is static. Ideally, the subject dataset would always be 'live' or be re-parsed every time the main AP+P website was updated, however, this is practically impossible when the process of creating the new data model is one of trial and error. The obvious solution is to improve the data model at the source, which is much easier said than done, as many cultural institutions are limited by the database structure of proprietary systems they use for maintaining catalogue records.

Having a very strong understanding of the data allowed me to make assumptions about how it might change over time. The greatest example of this occurs in *Decade Summary*. The quantity of prints in the collection from the early decades was relatively low, which meant I was able to include a container to view a larger artwork in the middle of the interface. This feature is based around a specific pattern in the data—a tricky thing to design for—but I can be reasonably confident that the number of early prints in the collection will not change dramatically.

API

The underlying structure of the data will determine, to an extent, what can be achieved with it. Another determining factor concerns how to actually access that data in a usable manner. Throughout my research project, I have relied on different API calls to provide access to the data. The API provides a connection between the MySQL database that holds the various tables of data from the site and returns pages of machine readable text, in the JavaScript Object Notation (JSON) format. Thanks to a funded research agreement between the NGA and UC we could commission the creation of these custom API calls. It is not feasible to load all the artwork data for every interface and so this approach allowed us to efficiently access specific subsets of the data and move some of the usual client-side processing to the server. For instance, we could access a subset of an artist's artworks or artworks by decade and print type, and then return the exact data we needed as JSON. This approach is evident in *Decade*

Summary: the initial API[1] call creates a summary of each decade on the server before it is returned to the client. This summary includes: a count of the total artists and their gender division; a count of the total artworks; the various media categories and print types for the decade and a count for each (an example is shown in Chapter five, page 109). To execute a script to create this summary in the browser would be quite intensive and take too long to load as it would require loading metadata for every single artwork.

Despite the power of pre-processing some of the data on the server, it will still require substantial transformation on the client-side. This is because each interface, and the views within it, require particular data structures that are different to the original output. For example, in *Works and Networks*, I load all the artwork data for the selected artist and then loop through the individual artworks in order to extract the creators for each one, I create a new data structure with this information and use it to produce the list of related artists in the left hand pane.

These interfaces are data-driven which means that it is critical to ensure the data structure is correct, as it directly influences the interfaces. As a result, regardless of the quality of the data, it will always need modifications in order for the interface to function effectively.

I use multiple API calls to load the data in the different interfaces, as it is impractical to load all of the data required at once. There are two main reasons for this: firstly, it would be an unreasonable request on the server (as it is intensive to return the API call); secondly, the user would spend some time waiting for the interface to load. Additionally, myriad different issues could occur whilst waiting for the page to load, from the server timing out or the user's internet connection being interrupted, all of which have implications in the data loading process. In *All Artists*, the data is loaded in blocks of 500 artists at time this allows me to show those first artists immediately, while waiting for the

1 http://printsandprintmaking.gov.au/dvapi/summary/decades/

remainder of the data to be returned. It is extremely important that the interface loads as fast as possible, as a slowly loading interface will lead to a diminished experience for the user.

The division between using jQuery (used in *Works and Networks*, *Decade Summary* and *All Artists*) and Angular (*The Fader*, *Subjects Explorer* and *Timeline*) is evident when dealing with the creation of a data-driven interface. Where JQuery is a JavaScript library that allows data manipulation, Angular is specifically based around the data model (MVC). The latter allows for two-way data binding, where any changes in the data model are updated immediately in the interface—one of its major benefits. To achieve the same effect in jQuery would require a number of additional functions. I will reflect on some of Angular's benefits later in this chapter.

Design language

In the following discussion I reflect on the development of a consistent design language across the six interfaces. This design language refers to the both the aesthetic appearance—use of typography, colours and layout; and the use of data-driven descriptive text within the interfaces. If the primary focus of an interface is to encourage exploration and discovery then the design must support that possibility occurring, being aware that elements of the design impact on how a user perceives the interface, its quality and its functionality.

Whilst producing the interfaces I was constantly aware of how they sat within the field of contemporary Web culture. This meant differentiating them from a traditional gallery website and showing the influences of modern Web design, but at the same time ensuring they did not become too experimental and confusing.

The typography of the interfaces is crucial because it determines how the interfaces are interpreted. Across the first four *Explore* interfaces I used the open-source Google Web Font 'Chivo', designed by Hector Gatti for Omnibus-Type foundry. Chivo is a sans-serif typeface with a slight futuristic feel and remains legible at smaller

sizes. At the time of their development in 2013, many websites were yet to embrace the possibilities of Web fonts and some, including the NGA, were still using the traditional 'Arial, Helvetica, sans-serif' font stack.[2] In the final two interfaces I used another Google Web font, Roboto, sans-serif font which is slightly more refined than Chivo. I chose Roboto because it had been implemented across many Google services and the Android operating system. I felt that it would ensure the interface would be less disorienting to users, as they would be more familiar with the typography.

In regards to colours employed across the various interfaces, I was keen to bring focus to the artworks themselves—rather than aspects of the interface. As a result, I use either white as the primary background colour (*Works and Networks*, *Decade Summary*, *Subjects Explorer* and *Timeline*) or light grey (*All Artists*, *The Fader*). In all interfaces except *The Fader*, I show the artworks with a white border around them. I found this provided enough contrast to make the artworks stand out from other aspects of the interface, and embraced Davies' (in Proctor, 2011) ideas about supporting the display of an artwork online, where the interface plays a similar role to the frame and so on. For *The Fader*, I decided the white border was not necessary as the other elements in the interface are so minimal that the focus is clearly placed on the large artwork images from the moment the interface loads. Where I did use additional colours, for instance in *Decade Summary*, I experimented with different combinations in order to create a combination that complemented the other aspects of the interface. In *All Artists* I decided to not use gender stereotyped colour combinations such as pink and blue (Cunningham and Macrae, 2011) and instead opted for a shade of orange for female, glacier blue for male, and light green for organisations. However, this did cause some confusion in the evaluation study, as participants did not immediately understand what genders the colours represented.

2 See for example, the NGA website in the Internet Archive Wayback Machine on April 9, 2013: https://web.archive.org/web/20130409040419/http://nga.gov.au/Home/Default.cfm. In the CSS: 'font-family': Arial, Helvetica, sans-serif;

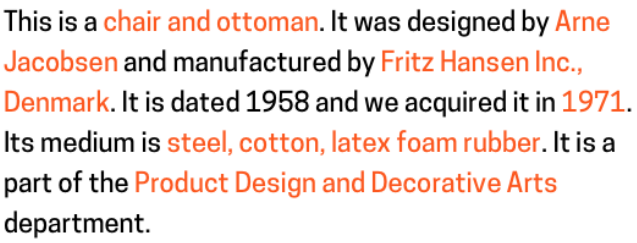


Figure 65
Detail of data descriptive text from Cooper Hewitt

I was also influenced by methods adopted by social media websites. For example, I examined how Facebook's timeline functioned whilst developing my *Timeline*; and online shopping sites such as ASOS, particularly through the use of facets when creating *All Artists* and *Subjects Explorer*. Researching these other sites was very useful as they are also trying to present large collections of data in the clearest, most engaging way— ensuring the user has the best possible experience—in order to maximise profits. Additionally, these sites often reflect current Web design trends seen in the use of varied Web fonts, multiple column layouts and Single Page Applications.

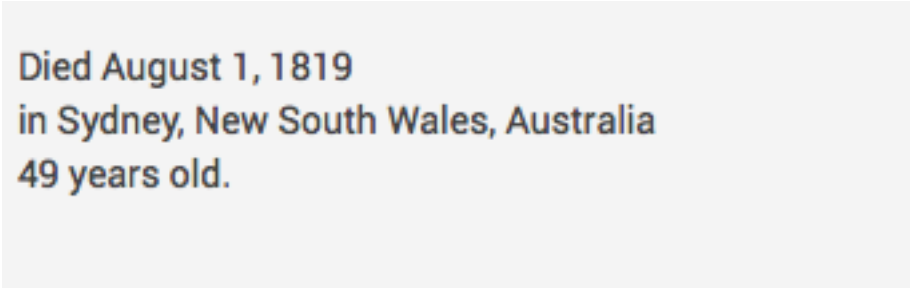
In the period between developing *Works and Networks*; *Decade Summary*; *All Artists* and *The Fader*, Google released their own visual design framework entitled, Material Design Guidelines.[3] I referred to the guidelines whilst developing *Subjects Explorer* and *Timeline*, as they included detailed methods for structuring the layout of various components, for example, the tabs in the *Subjects Explorer* summary are directly influenced by the Material Design tabs.[4]

There is no doubt that the aesthetic qualities of the interface can encourage and support free-form exploration and serendipitous discovery. However, there are also vital additional components that I see as being part of the design language developed through my research project. The first regards the use of data-driven text; the second is the importance of the permalink.

If the aim is to develop interfaces for the layperson, then the content within those interfaces must be displayed in the best possible way. This might be through a simple action such as reformatting the artist's name, to using data-driven descriptive

3 Google Material Design Guidelines: https://material.google.com
4 Material Design Tabs: https://material.google.com/components/tabs.html#

Figure 66
Detail of data descriptive text in *Timeline*



text as a feature in the interface. For instance, the format of an artist's name in the data is 'lastname, firstname', a structure clearly carried across from the library style catalogue system. However, there is no need for this format to remain in the interface and so in *Subjects Explorer* and *Timeline*, I reformatted the name to display as 'firstname lastname', which immediately makes the artists seem like real people, rather than mere names on a screen.

As noted in chapter 4, the use of data-driven descriptive text can be useful. In *Timeline*, if the data contains both the birth and death years then I can calculate the age of the artist when they died and display it in the interface. For example, J.W. Lewin[5] was born in 1770 and died in 1819, I can then show his age and where he died (Figure 65).

Erika Hall (2008) refers to this concept as "copy as the interface", and argues that language used on the Web should be simple, conversational, specific and direct. This is demonstrated clearly in the Cooper Hewitt collection site (Figure 66). I was able to create different ways to present the content and avoid using jargon or collection specific terms in order to create a better experience for the end user.

I have already discussed the role of the permalink in Chapter four, but it is necessary to reiterate its significance as a Web principle and key digital heritage concept. As Cameron (2008) has explained, the "networked object" can take on a life of its own, but a user always needs to be able to return to the original source if they want to.

5 J.W. Lewin Timeline: http://www.printsandprintmaking.gov.au/explore/timeline/#/artist/7900

Data Visualisation

The question, *"how can data visualisation techniques be developed to encourage exploration and discovery*?", framed my initial approach as I was primarily concerned with thinking about how data visualisation techniques could be used in the practical work. However, during the early production of *Works and Networks*, I realised that relying primarily on data visualisation techniques was not always the best option. The data, when visualised, can often be overwhelming and as such, a classic data visualisation approach can sometimes make the data harder to understand[6] (as illustrated in Figure 67). Rather than showing everything, it is important to provide the user with ways to start the exploration process. I believe that strong data visualisation has a time and a place[7], but rather than visualising for visualisation's sake (because it is the current trend), I considered it more worthwhile to investigate other options that offer more engaging methods of exploration.

In my work, I have borrowed techniques from data visualisation and combined them with modern Web design methods to create new and engaging interfaces that aim to encourage exploration and discovery. This has led to the creation of three new techniques for displaying rich cultural heritage data which I describe as: quantitative aggregation, dynamic focus+context displays and visual encoding.

> **Quantitative aggregation** refers to the use of word clouds and simple bar and column charts, to show counts and distribution of artworks. These provide a compact overview of the data and are a useful mode of navigation, for instance, through the facets in *All Artists*, the horizontal bar chart in and the decade bar chart in the summary of each subject in *Subjects Explorer*.
>
> In *All Artists* a number of facets allow filtering of the interface according to role, artwork count or gender. Rather than just being a simple checkbox, each facet includes a

6 See the DAAO for example: https://www.daao.org.au/visualisations/

7 For instance http://infovis.fh-potsdam.de/ddb/ and http://selfiecity.net

Figure 67
Experimental *Works and Networks* data visualisation showing relationships between artists

horizontal bar underneath the title. This is a bar-chart-like element that shows the patterns of distribution in the collection at a glance. Results from the evaluation study demonstrated that participants used these facets in the think-aloud observations, and 10% of the captured click events in the logged data were triggered by selecting a facet. The ability to filter was a technique users enjoyed (see P3:33). In addition to visualising this information the facets also allow refinement of the initial overview, allowing Shneiderman's 'zoom and filter' to occur.

Dynamic focus+context displays represent the creation of rich dynamic data-driven displays that provide detailed information about an artwork without losing overall context. These occur in most of the interfaces, although are particularly successful in *Subjects Explorer* and *Timeline*, where the use of Angular made their creation much more efficient. In *Subjects Explorer*, a dynamic focus+context display is shown when a user clicks on an artwork thumbnail and the full width row is inserted into the display; in *Timeline*, the action is similar, larger artwork images alongside the thumbnails, and detailed exhibition information is inserted into the interface. This technique is informed by a similar style of display in Google Image search, however it is unique in this situation as it used as a platform for exploration.

Visual encoding describes a technique where a data feature is encoded in a single data element, such as each artist being represented with a colour band for gender in *All Artists* or the artwork count being reflected in the font size in *Subjects Explorer* and *Timeline*. These encodings function differently to quantitative aggregation as they render data at the individual element level and integrate the representation with standard Web content (such as font size or width); whereas, quantitative aggregation focuses on providing a compact overview using standard

visualisation techniques. A single data feature, for example, the artwork count, can be represented using both quantitative aggregation and visual encoding techniques.

The use of bands of colour in *All Artists* was received positively in the evaluation study and I consequently implemented a similar technique in *Subjects Explorer* and *Timeline*. Because participants took a few seconds to understand exactly what the colours referred to and to avoid any potential confusion regarding their meaning, I provided a key: in *Subjects Explorer* brief instructions mimic the display of the subject terms and in *Timeline* the various entities are colour coded in the header.

In *All Artists* I adjusted the width of the artist box so it reflects their artwork count; in the summary information for a subject in *Subjects Explorer*, the font size is determined by how many artworks it represents; a similar font size adjustment is made in *All Artists*. These visual signals reinforce aspects of the data, either quantitative or conceptual (gender, entity type). The strength of these techniques in encouraging a user to interact with the dominant element was shown in the logged data for *All Artists*, where the second highest rate of recorded events was on the first of the larger artist boxes in the interface. It seems that adjusting the size of the various elements can encourage the user to 'read' the interface and potentially start to interpret it, before clicking on an item of interest. This is a simple, but powerful technique as it immediately allows the user to scan through the display and quickly understand the distribution of that element in the interface as well as providing a prompt for interaction.

These techniques are novel in this application but are applicable to any other collection of cultural data and represent unique contributions to the field of knowledge.

In *Works and Networks* I show the relations between artworks and creators with interactive highlighting. This technique works reasonably well in terms of visualising connections which would otherwise remain hidden in the data. An alternative approach visualising this data using network graphs had shown that it was impossible to read. Feedback from participants in the evaluation study suggested that interactive highlighting in *Works and Networks* worked well in revealing the connections but could be frustrating because the user cannot scroll through the artworks at the same time as hovering over the artist name.

However, for my work it is important to not rely on a classic data visualisation technique (such as a network graph) as the only method to visualise the information. I have found it more productive to regard data visualisation techniques as one tool among many for encouraging free-form exploration and serendipitous discovery.

New Web-based technology

The discussion that follows addresses the question *"how can new Web based technologies be leveraged to create exploratory digital interfaces?"* I refer to three concepts that informed the development of my practice-based work. Firstly, I will discuss the experimental approach that provided a foundation for the project before reflecting on Single Page Applications and the use of Client Side processing.

Experimental Approach

Throughout my research project I adopted an experimental approach which involved embracing new techniques and methods and utilising them in many different ways. For example, I trialled the adoption of various frameworks—from Angular to Bootstrap; developed new methods to handle the display of images; and made the interface accessible online for 'live testing'.

Jeff Croft (2007) defines the framework as "a set of tools, libraries, conventions, and best practices that attempt to abstract routine tasks into generic modules that can be reused." He explains that the goal is to allow the designer or developer to focus on tasks

that are unique to the project, rather than reinventing the wheel each time around (Croft, 2007). The discussion that follows will reflect on the use of two different frameworks, Angular and Bootstrap, that I regard as crucial to the success of the final interfaces.

As noted in Chapter four, I used the Angular Web development framework in the development of *The Fader*, *Subjects Explorer* and *Timeline*. This proved to be an extremely worthwhile decision because Angular perfectly aligned with the dynamic data I was working with. It contains a number of directives (HTML attributes whose functionality is extended upon in Angular) that allowed me to commence sketching working prototypes much faster than in the development of the first three interfaces. From my perspective, one of the greatest features of Angular is the separation of presentation, content (data) and logic from each other. In this regard, the HTML should contain the content, the CSS should be used only for presentation and the logic should be contained in the JavaScript files.

Unfortunately, when using jQuery, these three components are mixed together, for example, in *Works and Networks*, we insert all of the content into the HTML page after it has been loaded. Whilst this results in a smaller HTML file—there are only 16 lines of code that actually display the interface—it means there is a substantial amount of HTML in the JavaScript file, the consequence being that it becomes difficult to work out where different elements originate.

However, in stark contrast, Angular allows the use of expressions in the HTML pages which will be updated automatically when the data model changes. This means that the HTML page contains all of the tags used to display the interface, while the JavaScript file only contains the logic. Once a function runs in the Angular JavaScript file it will update the data model, which will then update the interface. For example, in *Subjects Explorer*, when viewing a subject and its artworks, a summary of the top artists, print types and media categories is shown. When a user selects one of the names, I simply replace the current artworks data with the selected artist's artworks data, and the display

updates immediately. It is an extremely powerful technique.

Angular challenged me—I had to learn another programming language and think differently about the logic behind some of the programming choices. However, I now consider Angular to be superior to jQuery and there is no doubt that I will use it for future work in this area.

On the other hand, the adoption of Bootstrap[8] was a less successful endeavour. Bootstrap is an open-source framework originally created by a development team at Twitter. It contains a large number of components designed to enable faster front-end Web development; these range from simple base typographic styles to a responsive grid system. This system is based on a 12 column grid which allowed me to use a number of predefined classes to specify how wide the various elements in the interface were and how they would appear on different screen sizes. For example, in *Timeline*, the exhibitions and references in the left pane have the classes "col-md-3" and "col-xs-12", the first class specifies the content to be three columns wide on medium screens; the second class will only be applied on extra small screens (mobiles) and specifies the width to be 12 columns wide (full screen). Despite these seemingly time-saving components, I found that, I spent a vast amount of time overriding the default styles with my own. The appearance of the interface was strongly influenced by the possibilities of Bootstrap. In hindsight, it would have been more effective to simply write all of my own CSS from scratch or use a more customisable CSS-grid framework like Skelton[9] or Susy[10], as I had with the previous interfaces, rather than writing CSS to override many of the default Bootstrap styles.

In regards to artwork images, I was very aware that the sheer quantity being displayed in the interfaces could have huge implications on the page load time. I therefore experimented with different ways of loading and preparing them for display. As already noted, I retrieve the images directly from the NGA server, rather

8 Bootstrap: http://getbootstrap.com
9 http://getskeleton.com
10 http://susy.oddbird.net

than the server that the AP+P website is hosted on. There are two reasons for this, firstly, loading the images directly from the NGA server was faster (due in part to the location of their server within a high speed data centre in Canberra); and secondly, because new images are constantly being added to the NGA website and there is a delay before they appear on the AP+P website. This approach ensures that I can display as many images as possible in the interfaces.

In all of the interfaces, I load each artwork image as a CSS background image—rather than a HTML image—because of the complexities of aligning an image vertically and horizontally within a grid. This technique allows me to check if the API data contains an image; if it doesn't then I can immediately replace it with the 'no-image' graphic, or in the *Subjects Explorer* and *Timeline*, to display the title, date and 'no image available' text. It is not possible to replicate the same technique if the image is placed with a HTML tag.

In order to reduce page load time in *Decade Summary*, the images are not cropped—instead we load the smallest size from the NGA server (135x90 pixels), set a width and height in the CSS (42x42 pixels) and hide the overflow—this has the effect of cropping a tight square in the middle of the image. Then when a user hovers over the cropped artwork, we simply remove the constrained width and height and show the full-size image, as it has already been loaded it is displayed immediately.

Once I decided the interfaces were near completion I would place them online in a 'beta'-like testing approach that I refer to as 'live testing'. I found there to be real value in publishing the interfaces online in their 'beta' state, rather than waiting until they were completely finalised. This approach allows for bug testing to be carried out by the end user, with feedback received by email and through social media. For instance, after uploading *Subjects Explorer* I received valuable feedback from a user on Twitter about an uncaught error in the subject summary. It is not possible for a single person to test every different aspect of the interface when it is built with dynamic data and so feedback such as this is invaluable.

Single Page Applications

As noted in Chapter four, the interfaces are considered to be Single Page Applications (SPAs). The idea being that the interface functions more like a native application, one where the content or view updates immediately upon interaction; rather than a traditional website where the page would have to be reloaded. The single page approach allows for the development of fast, complex applications that feel more solid and responsive (due to the speed of the page) than traditional websites (Penman, 2015). However, by only ever having one page, a traditional HTML link can no longer exist. I overcame this by appending hash fragments to the URL in *Works and Networks*, *Decade Summary*, *Subjects Explorer* and *Timeline*. This technique allowed the current view to be bookmarked or shared but there were limitations. For example, in *Works and Networks*, participants were annoyed when they clicked the back button and the interface didn't remember the last artwork they were looking at. This was because I appended the artist id to the URL, but not the individual artwork ids—which would need to be constantly updated. Unfortunately, for this technique to work, it was not simply a matter of adding the relevant information to the URL, it also required additional functions to interpret the URL and load the correct data.

Additionally, the SPA model makes the implementation of Google Analytics less straightforward than usual. Instead of using the traditional tracking code, I had to create a number of custom Google Analytic events that were added to various points of interaction within the interfaces. This allowed me to track interactions within the interface as if they were individual pages. In the development of SPAs this is an important aspect to be aware of as Google Analytics is the tracking tool of choice within many cultural institutions and the metrics it captures are often used to evaluate success of the project (Finnis et al., 2011).

Client Side processing

One of the methods I identified in Chapter three was the use of the Web. I stated that I was interested in experimenting with the

possibilities of client-side processing, rather than using server-side processing as would have been the case in a more traditional website. This approach led to some challenging issues which I had to overcome to ensure the interfaces worked successfully.

There is no doubt that the interfaces must be fast— they need to load quickly and respond to interaction immediately— a given within modern Web development. The speed of the interface was mentioned a number of times in the Web survey, for example, Participant 175 wrote: "I truly appreciate that however you implemented this, it is not real time intensive, and does not bog down my browser." This feedback showed that it was important to not overload the browser and cause it to slow down dramatically or crash. The biggest issue in regards to page speed is how fast the data itself can be loaded and processed. Limitations are caused by the speed of the user's connection and by how quickly the browser can process that data. In order to overcome this I ensure that the size of the data being downloaded is as small as possible and that any client-side processing is not too intense. I was surprised at the efficiency with which the browser was able to process large amounts of data on the client-side, specifically in *All Artists*, where it quickly processes user interaction and refines the display accordingly. Over the course of this research project there has been a shift towards client-side processing, as witnessed by the development of new JavaScript libraries such as PourOver[11] and Tamper[12], specifically built to help sort and filter large collections in the browser (Koski and Hinton, 2014). Whilst I do not use these in my own work, it shows the appropriateness of my own approach.

As I have already discussed, the majority of data used in the interfaces comes from different API calls that are processed on the server, before returning JSON which is then modified on the client-side and used to produce the interface. In the case of *Subjects Explorer*, I created static JSON files for each subject which

11 PourOver— a library for fast filtering and sorting of large collections in the browser: http://nytimes.github.io/pourover/

12 Tamper—a serialization protocol for categorical data: http://nytimes.github.io/tamper/

are stored directly on the server, with only one file being loaded at a time. These JSON files are then transferred from the server with GZip[13] compression which dramatically reduces the file size, for example, the uncompressed ppm-subjects.json is 1.2MB and only 244KB when transferred; the largest file: australia.json is reduced from 1.5MB to only 190KB. In *Timeline* I combined a static JSON file with live API calls to ensure the interface loaded as quickly as possible. An additional PHP script on the server updates the static list of artists every 24 hours to ensure it reflects the current state of the data. These different processes of loading the data have all been created to ensure the interface loads as quickly as possible for the user, whilst also limiting the load on the server itself.

7.3 Exploration

In regards to the broad theme of exploration, I will reflect on the concept of overview, before considering the role of exploration and visualisation, and finally discussing the use of evaluation.

Overview

I have been strongly influenced by the concept of the overview throughout my project. It forms the first stage of Shneiderman's (1996) Visual Information Seeking mantra and was a key feature in many of the websites discussed in the Practice review in Chapter two. Of the six interfaces, only *The Fader* does not start with an immediate overview.

Production of the interfaces has shown that creating an effective overview can be difficult. The challenge is to represent the whole collection in a compact form whilst still being usable, however, the more data that exists, the harder it is to display. In *Works and Networks*, I resolved this by choosing to provide a predefined starting point— essentially narrowing the initial overview so that, rather than representing all artworks in the collection, it provides a focused overview—Whitelaw (2015) calls

13 GZip: http://www.gzip.org

it an 'innerview'—of all the artworks from the selected artist. The concept of an innerview is useful as it allows me to provide a complementary view to the other interfaces. For example, *Decade Summary* and *All Artists* provide comprehensive overviews whereas *Works and Networks* provides a more immersive innerview.

Shneiderman (1996); Greene et al. (2000); Keller and Tergan (2005) all argue that overviews are an effective technique for revealing the size and the diversity of a collection. I understood that by adding interaction (as demonstrated by Ferreira de Oliveira and Levkowitz, 2003), an overview could become a valuable technique for encouraging information-seeking to occur within a dynamic data-driven interface such as the ones I have developed. The results of the evaluation study (see section 6.2) show that interfaces containing interactive overviews facilitate a process of free-form exploration and serendipitous discovery. The open-ended nature of the interfaces did not pose any substantial problems, as users were willing to engage with the interfaces and sought to understand how they worked, or what they were trying to reveal. It is an understanding of the user that aligns with Dörk's (2011) concept of the information flaneur and one that shows that users are prepared to explore if they are provided with the interfaces that make it possible.

Exploration and visualisation

In my reflections on data visualisation earlier in this chapter, I argued that relying primarily on visualisation techniques is not always the best option, and the application of its techniques should be considered within different contexts, specifically in combination with Web design methods. My discussion of the overview emphasised its role as the starting point of an effective data visualisation. Furthermore, feedback showed the effectiveness of the overview as a useful starting point; however, an interesting division emerges. When creating *Subjects Explorer* and *Timeline* I found myself asking: *is it more useful to show everything? Or is it better to show 'enough' to encourage the user to start exploring?*

These questions have led me to identify a distinction between the models of visualisation and exploration.

A visualisation model is based upon many underlying assumptions, for instance, we assume there is a dataset that can be shown and that an overview, or a 'show everything' approach, will provide a useful starting point. Often, these representations are practically static, and only include basic dynamic focusing and filtering, as described in Ahlberg and Shneiderman's early work in the practice review (see section 3.4).

On the other hand, exploration is a process into which a user can be invited. Through the development of rich hooks, I can provide multiple opportunities for the user to begin the exploration process, as I will demonstrate with reference to initial views of both *Subjects Explorer* and *Timeline*.

In *Subjects Explorer*, the view includes the full list of subjects in the left pane, and in the right, the top subjects are repeated, with the font size relative to the artwork count—giving the user two different ways to start exploring. In *Timeline*, a random sample of artists is shown, again using the font size adjustment technique, which aims to prompt the user to make a selection. These are rich hooks. They do not attempt to show everything or provide an overview but instead focus on showing 'enough' to encourage the user to start exploring the collection.

Where visualisation defaults to a static view of the data, exploration is about constantly moving through it. This can be achieved through rich hooks, links (between artworks and between the individual interfaces) and facets in the data (for instance, through the checkboxes in *All Artists*). Providing these different methods can allow the user to refine the data exploration.

I have shown that my interfaces do not rely on the visualisation model and have described the need for more substantial models of representation that use rich hooks to encourage exploration, rather than relying on static representations.

Generous interfaces

In the literature review, I described how Whitelaw's (2012) concept of generous interfaces was directly informed by the production of *Works and Networks*, *Decade Summary* and *All Artists*. The principles provide a useful framework to consider when developing this style of access. In order to extend understanding of these principles I illustrate how they are manifested in my work.

> I always **show the data first**—present multiple ways for the user to start exploring the collection—rather than starting with a search box. This might be through detailed representations of the data, as shown in *Decade Summary* or through the random sampling of artists at the beginning of *Timeline*. In *The Fader* a large artwork image is shown immediately and the user can then decide if they wish to change focus.
>
> Five of the interfaces **provide rich interactive overviews** that allow the user to orient themselves within the collection and find a starting point for exploration. Feedback collected through the evaluation study shows that they are effective in encouraging the exploration process to occur.
>
> It can be hard to display an entire digital collection, therefore in *Decade Summary*, I **provide samples** of the data—through the cropped thumbnails—which provides a condensed visual representation of the collection. The display of the data is conceived as a way of providing contextual clues that encourage exploration.
>
> In all the interfaces I allow a user to explore the collection whilst also **providing context**. This is primarily achieved through the creation of dynamic focus+context displays, for example, in the three-pane display in *Works and Networks* or the artwork and exhibition view in *Timeline*. The focus+context style of display aligns with the latter stages of Shneiderman's Visual Information Seeking mantra, namely "zoom and filter, details on demand"; in

my work I combine all these stages into a single display rather than splitting them across multiple views.

In all the interfaces I **share high quality primary content**. I do this by always providing a permanent link back to the original networked object on the AP+P website. It became apparent in the evaluation study that users followed these paths to the catalogue reference, however, the process of accessing them needs to be consistent, to avoid the problem of some links not opening in a new tab or window.

In addition to these principles, I suggest that an additional two could be added to the list: **use less jargon** and **be more descriptive**. As shown previously in this chapter, the creation of data-driven descriptive text is a useful means of ensuring that content is displayed in a user friendly way. This aligns with the "descriptive cataloguing" approach undertaken by gallery staff when cataloguing artworks for display on the AP+P website (Butler, 2013) and Hall's (2008) "copy as the interface."

The best way to stimulate curious, creative and critical engagement is to create interfaces that allow multiple methods for the user to start exploring, as well as providing detailed representations of the data within. *Decade Summary* is a playful interface using dual overviews that entice interaction—it is through interaction that detailed insights into the data will be revealed. *Subjects Explorer* allows a curious user to investigate how subjects are related to artworks and vice versa; again with detailed data for each artwork being presented within the single view. On the other hand, I consider *Timeline* to be targeted at a more informed user—one who has already played with the other interfaces, or might already have an understanding of the data. It allows the user to select a random name, or search for an artist, and then view comprehensive information, all displayed within a single interface.

Evaluation

In order to validate the effectiveness of the creative works and determine how the interfaces were used I completed a mixed-method evaluation study. From the outset I hypothesised that the open-ended nature of the exploration process would make the evaluation of such work difficult. In hindsight, I realised that the problem was not so much the free-form exploration, but choosing the most appropriate set of methods for the evaluation study. The combination of think-aloud observations, a Web survey and extensive data logging provided me with a considerable amount of data which provided an illuminating insight into how the interfaces were used and the aspects which users did and did not enjoy.

The think-aloud observations provided the most revealing feedback. This was because I could actually watch the participant interact with the various interfaces in real time; seeing where they were annoyed and what they enjoyed. They provide qualitative evidence of the 'delight' factor, which would never be experienced from a search interface. This was in contrast to the Web survey, where it was often unclear what the respondent was referring to or what an error was caused by. For example, Participant 190 wrote that they did not enjoy "poor website performance" and also noted that they "did not load", but the server logs do not show any server downtime, which suggests the problem may have been at the participant's end. However, by not being able to see the participant's screen it is hard to determine the exact cause of the problem they experienced.

My dual role as developer of the interfaces and observer in the think-aloud observations placed me in a unique position where I gained many insights into how the participant used the interfaces. I would not have been able to gain the same understandings if someone else completed the observations.

The format of the mixed-method evaluation study proved to be particularly effective in allowing me to evaluate the success of interfaces that encourage free-form exploration and serendipitous discovery, as it provided me with substantial

qualitative and quantitative data. I then used these results to inform the development of the final two interfaces.

7.4 Contributions

My key contribution to new knowledge in the field of cultural heritage and design is through the creation of six interfaces which allow the visual exploration of a large digital cultural heritage collection. These interfaces are unique in the field and focus on providing novel modes of collection access that encourage free-form exploration and serendipitous discovery. This has been achieved through the development of new techniques for promoting exploration within a Web-based context. Specifically, I refer to quantitative aggregation, dynamic focus+context displays and visual encoding. These techniques combine methods from data visualisation (such as Shneiderman's Visual Information Seeking mantra and focus+context displays) with those from modern Web design (including HTML5, CSS3 and Angular) and enable exploratory modes of information seeking to occur. These techniques are transferable to all digital collections. Through the evaluation of the interfaces, I have demonstrated how mixed-method research studies can provide valuable insights into this style of practice-based work.

7.5 Conclusion

In this chapter I have reflected on the interfaces I produced over the course of my research project and outlined my contributions to new knowledge. I addressed my research questions by reflecting on the themes of technical development and exploration. Regarding technical development, I have shown the sketching in code methodology is appropriate for this style of work. It has also been demonstrated that the possibilities of what can be produced are bound to the strength of the data. As such, I have described how structured data is critical, and that using multiple API calls is the most effective way to access it. My reflection on the development of a consistent design language shows the importance of the underlying

aesthetic qualities of the interface, and also suggests that using data-driven text can allow the user to experience the content in more engaging ways. I have demonstrated furthermore that relying primarily on data visualisation techniques has limitations, which I addressed by combining data visualisation and Web design methods in order to create new techniques—quantitative aggregation, dynamic focus+context displays and visual encoding—that are well suited to displaying rich cultural heritage data. I have shown how the use of new Web-based technology has proven to be effective in the production of my creative works, as the experimental approach allowed me to embrace new techniques and methods as the project progressed. This led to the extremely worthwhile adoption of the Angular framework.

In the second part of the chapter I reflected on the theme of exploration and discussed the concept of the overview. I argued that overviews can be effective but also have limitations, I suggested that they should be considered as one potential method for encouraging exploration and discovery in digital cultural heritage collections. In order to consider other methods, I distinguished between the models of visualisation and exploration, and outlined how the development of rich hooks can entice users into the exploration process. Following on from this, I illustrated how the principles of generous interfaces were demonstrated in my work and I suggested that additional principles—use less jargon and be more descriptive—could be useful additions. I have shown how I utilised an evaluation study as a formative component of my research project and discussed how the mixed-method format was appropriate, given the interfaces aim to encourage free-form exploration and serendipitous discovery. Finally, I clearly outline my contributions to new knowledge in the fields of cultural heritage and design.

In the following chapter, I will summarise the project, its overall contributions to knowledge and discuss how the techniques I have developed are applicable to any digital collection.

8 Conclusion

When I began my PhD in August 2011, galleries and museums were becoming increasingly concerned with expanding public access to their rich collections. These concerns were framed in response to meeting government funding requirements, an increasingly competitive cultural sector and the need to secure larger audience numbers (Davis and Howard, 2013). The Web provided unprecedented new access opportunities but most galleries and museums did not initially exploit its possibilities. Instead, they placed their limited catalogue records online and relied on keyword-based search as the primary method of access. There were however some leaders in the field who experimented with innovative new methods for collection access in the online space. As outlined in my practice review they included SFMOMA, with their exploration-focused *ArtScope* tool, and more recent websites by the Walker Art Centre (2011), Rijksmuseum (2012) and Cooper Hewitt (2016) which emphasised the visual quality of the collections they provided access to, rather than requiring the user to search the collection first. These examples provide users with immediate access to the collection data and display it in dynamic, engaging ways which represent a significant departure from the norm.

In my work, initially undertaken as in collaboration with Dr Mitchell Whitelaw and supported by a grant from the NGA and UC, I focused on using data from the AP+P collection to create a number of visual information seeking interfaces that aimed to encourage free-form exploration and serendipitous discovery. By adopting a practice-based approach, I sought to pursue new directions in online access to cultural collections that were driven by the concept of generosity.

The first three interfaces, *Works and Networks*; *Decade Summary* and *All Artists*, were produced in collaboration with Whitelaw. They formed the foundation of the research project and applied the key concepts that had emerged through my literature review. They tested the possibilities of the overview as outlined

in Shneiderman's Visual Information Seeking mantra, as well as providing different ways for visual information seeking to occur. The resulting interfaces were then critically evaluated in 2014 through a mixed-method evaluation study that tested their effectiveness in stimulating free-form exploration and serendipitous discovery. The results from the evaluation study indicate that users are willing to explore a digital collection if they are provided with interfaces that open up their options for free-form exploration to occur.

I subsequently created a further three works, *The Fader*; *Subjects Explorer* and *Timeline*, the latter two of which were directly informed by the results of the evaluation study I conducted. I consider the final two interfaces, *Subjects Explorer* and *Timeline*, to be the most complex and substantial of all the works produced, as they responded to the feedback from the evaluation study and also encapsulate the key concepts investigated throughout the project.

The interfaces are not mere prototypes with an indeterminate future. They are working tools that were launched online and are used every day by specialists and the general public around the world. They have been discussed in various international forums, including the leading industry conference *Museums and the Web* held in Portland, Oregon in 2013 and *Museums and the Web Australia* in Melbourne in 2015. The project achieved 'First Runner Up' in the *Best DH Data Visualization* category in the prestigious *Digital Humanities Awards* in 2015, a further indication of its positive reception.

During my candidature there have been significant technical developments in the various fields in which my research is located. Wherever possible I have incorporated them into my work through an adaptive experimental approach that involved the use of different techniques, the most significant of which was the adoption of the Angular framework. Angular's inbuilt processes perfectly aligned with the dynamic data models I utilised in the project and allowed me to develop more interesting and engaging interfaces

In the practice and literature review I have outlined why galleries and museums need more engaging methods of online collection access. The six interfaces I developed are dynamic and data-driven. They combine methods from data visualisation and modern Web design to create new techniques to display cultural heritage data and foster exploratory modes of information seeking. In emphasising the visual and offering multiple ways to access a collection they make an ongoing contribution to a field which is in a rapid and exciting phase of development. The methods, techniques and overall approach I have developed for the AP+P site are applicable to any other collection of cultural data.

My research has contributed new knowledge about how to create interfaces which encourage free-form exploration and serendipitous discovery in digital cultural heritage collections. I have achieved this through the production, publication and evaluation of the practice-based body of work. What I have created represents an investigation into the possibilities afforded by applying Web design and data visualisation techniques to richly structured cultural heritage data. The production process brings novel insights into how abstract concepts, such as Shneiderman's Visual Information Seeking mantra, can be utilised within the specific technical context of Web-based exploratory interfaces. Each of my interfaces provides an innovative method of accessing the online AP+P collection, maintaining context without sacrificing detailed content—in contrast to traditional online collection interfaces. The context for the works I have created recognises that alternatives are required to encourage and facilitate exploratory or serendipitous information seeking experiences that are more in keeping with contemporary Web usage. Through the mixed-method evaluation study, I have contributed further insights about how to evaluate open-ended exploratory interfaces and provided useful reflections for future work. My research project has shown that it is possible to escape from the constraints of the search box and create interfaces which encourage exploration and discovery in digital cultural heritage collections.

9 Bibliography

Ahlberg, C., & Shneiderman, B. (1994). Visual information seeking: tight coupling of dynamic query filters with starfield displays (pp. 313–317). ACM Press. http://doi.org/10.1145/191666.191775

Ahlberg, C., Williamson, C., & Shneiderman, B. (1992). Dynamic Queries for Information Exploration: An Implementation and Evaluation (pp. 619–626). ACM Press.

Armitage, T. (2009). Toiling in the data-mines: what data exploration feels like. Retrieved from http://berglondon.com/blog/2009/10/23/toiling-in-the-data-mines-what-data-exploration-feels-like/

Armitage, T. (2015). Sketching and Engineering. Retrieved from https://goodformandspectacle.wordpress.com/2015/01/12/sketching-and-engineering/

Asami, S., & Patterson, D. (2000). *GridPix – A Method for Presenting Large Image Files Over the Internet* (Computer Science Division (EECS) No. UCB/CSD-00-1099). Berkeley, California: University of California.

Backhausen, D. T. J. (2012). Adaptive IR for exploratory search support. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 992–992). New York, NY, USA: ACM. http://doi.org/10.1145/2348283.2348416

Bailey, J., & Pregill, L. (2014). Speak to the Eyes: The History and Practice of Information Visualization. *Art Documentation: Journal of the Art Libraries Society of North America*, *33*(2), 168–191. http://doi.org/10.1086/678525

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, *13*(5), 407–424.

Bates, M. J. (2007). What is browsing—really? A model drawing from behavioural science research [text]. Retrieved 20 June 2016, from http://www.informationr.net/ir/12-4/paper330.html

Bearman, D., & Trant, J. (1999). Interactivity comes of age: museums and the World Wide Web. *Museum International*, *51*(4), 20–24. http://doi.org/10.1111/1468-0033.00225

Belkin, N. J. (1993). Interaction with texts: Information retrieval as information-seeking behavior. In *Universitätsverlag Konstanz* (pp. 55–66).

Bell, D. (2009). On the net: navigating the World Wide Web. In G. Creeber & R. Martin (Eds.), *Digital Culture: Understanding New Media* (pp. 30–46). Berkshire, UK: McGraw-Hill Education.

Berners-Lee, T. (1998). Frequently Asked Questions. Retrieved from https://www.w3.org/People/Berners-Lee/FAQ.html

Bertini, E. (2011). Can visualization influence people? I mean can we prove it? Retrieved 21 June 2016, from http://fellinlovewithdata.com/reflections/can-visualization-influence-people-i-mean-can-we-prove-it

Borgman, C. L., Hirsh, R. G., Walter, V. A., & Gallagher, A. L. (1995). Children's searching behavior on browsing and keyword online catalogs: the science library catalog project. *Journal of the American Society for Information Science*, *46*, 663–684.

Bowen, J. (1999). Only connect. *Museum International*, *51*(4), 4–7. http://doi.org/10.1111/1468-0033.00222

Bowen, J., Bennett, J., & Johnson, J. (1998). Virtual Visits to Virtual Museums. Retrieved from http://www.archimuse.com/mw98/papers/bowen/bowen_paper.html

Bowen, J. P. (2010). A Brief History of Early Museums Online. *The Rutherford Journal* (*The New Zealand Journal for the History and Philosophy of Science and Technology*), *3*. Retrieved from http://www.rutherfordjournal.org/article030103.html

Brenner, S. (2015). Reconsidering searching and browsing on the Cooper Hewitt's Collections website. Presented at the Museums and the Web 2015, Chicago. Retrieved from http://mw2015.museumsandtheweb.com/paper/reconsidering-searching-and-browsing-on-the-cooper-hewitts-collections-website/

Bruce, H. (2005). Personal, Anticipated Information Need. *Information Research: An International Electronic Journal*, *10*(3). Retrieved from http://eric.ed.gov/?id=EJ1082030

Burton Jones, K. (2012). The Transformation of the Digital Museum. In P. F. Marty & K. Burton Jones (Eds.), *Museum Informatics: People, Information, and Technology in Museums* (pp. 9–25). Routledge.

Butler, R. (2011). Prints and printmaking: web innovation and excellence. Retrieved from http://printsandprintmaking.gov.au/about/

Butler, R. (2013). Prints and printmaking: web innovation and excellence. Retrieved from http://printsandprintmaking.gov.au/about/prints-printmaking-web-innovation-excellence/

Cailliau, R. (2000). A little history of the World Wide Web. Retrieved from http://www.w3.org/History.html

Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web Survey Methodology*. SAGE.

Calore, M. (2010, April 22). April 22, 1993: Mosaic Browser Lights Up Web With Color, Creativity. Retrieved from http://www.wired.com/2010/04/0422mosaic-web-browser/

Cameron, F. R. (2008). Object-oriented democracies: conceptualising museum collections in networks. *Museum Management and Curatorship*, *23*(3), 229–243. http://doi.org/10.1080/09647770802233807

Candy, L. (2006). *Practice Based Research: A Guide* (Creativity and Cognition Studios Report No. V1.0). Sydney, Australia: University of Technology Sydney. Retrieved from http://www.creativityandcognition.com/content/category/10/56/131/

Candy, L. (2009). *Practice-Based Research and the PhD: A Study*. Sydney, Australia: University of Technology Sydney. Retrieved from http://research.it.uts.edu.au/creative/linda/PBR/

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann.

Case, D. O. (2012). *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior*. Emerald Group Publishing.

CERN. (2003). Ten Years Public Domain for the Original Web Software. Retrieved from http://tenyears-www.web.cern.ch/tenyears-www/Declaration/Page1.html

Chang, Y.-W. (2013). The influence of Taylor's paper, Question-Negotiation and Information-Seeking in Libraries. *Information Processing & Management*, *49*(5), 983–994. http://doi.org/10.1016/j.ipm.2013.03.003

Chan, S. (2007). Tagging and Searching – Serendipity and museum collection databases. Presented at the Museums and the Web, Toronto. Retrieved from http://www.museumsandtheweb.com/mw2007/papers/chan/chan.html

Chan, S. (2011). The museum website as a newspaper – an interview with Walker Art Center. Retrieved from http://www.freshandnew.org/2011/12/museum-website-newspaper-interview-walker-art-center/

Chan, S. (2012). First look at our new online collection (public alpha). Retrieved 18 June 2016, from http://labs.cooperhewitt.org/2012/online-collection-alpha/

Circus Oz. (2012). The Living Archive. Retrieved 6 June 2015, from http://archive.circusoz.com

Clarke, A. (2008). Time to stop showing clients static design visuals. Retrieved from https://stuffandnonsense.co.uk/blog/about/time_to_stop_showing_clients_static_design_visuals/

Coburn, J. (2016). I don't know what I'm looking for: Better understanding public usage and behaviours with Tyne & Wear Archives & Museums online collections | MW2016: Museums and the Web 2016. In *MW2016*. Retrieved from http://mw2016.museumsandtheweb.com/paper/i-dont-know-what-im-looking-for-better-understanding-public-usage-and-behaviours-with-tyne-wear-archives-museums-online-collections/

Cooper Hewitt. (2016). Cooper Hewitt Collection. Retrieved 6 June 2015, from https://collection.cooperhewitt.org

Cooper Hewitt. (2014). About Cooper Hewitt. Retrieved 22 June 2016, from http://www.cooperhewitt.org/about/

Cope, A. (2013). 'B' is for beta. Retrieved 22 June 2016, from http://labs.cooperhewitt.org/2013/b-is-for-beta/

Couper, M. P. (2000). Review: Web Surveys: A Review of Issues and Approaches. *The Public Opinion Quarterly*, *64*(4), 464–494.

Coyier, C. (2013). Job Titles in the Web Industry. Retrieved 19 May 2016, from https://css-tricks.com/job-titles-in-the-web-industry/

Craft, B., & Cairns, P. (2009). Sketching Sketching: Outlines of a Collaborative Design Method. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology* (pp. 65–72). Swinton, UK, UK: British Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=1671011.1671019

Croft, J. (2007, June 12). Frameworks for Designers. Retrieved 2 June 2016, from http://alistapart.com/article/frameworksfordesigners

Cunliffe, D., Kritou, E., & Tudhope, D. (2001). Usability Evaluation for Museum Web Sites. *Museum Management and Curatorship*, *19*(3), 229–252. http://doi.org/10.1080/09647770100201903

Cunningham, S. J., & Macrae, C. N. (2011). The colour of gender stereotyping. *British Journal of Psychology*, *102*(3), 598–614. http://doi.org/10.1111/j.2044-8295.2011.02023.x

Davis, W., & Howard, K. (2013). Cultural policy and Australia's national cultural heritage: issues and challenges in the GLAM landscape. *The Australian Library Journal*, *62*(1), 15–26. http://doi.org/10.1080/00049670.2013.774684

DePaolo, C., Szukala, Z., & Kendrick, E. (2012). Building an Online Commons for the Park. Presented at the Museums and the Web 2012, San Diego, California. Retrieved from http://www.museumsandtheweb.com/mw2012/papers/building_an_online_commons_for_the_park.html

DHA. (2015). Digital Humanities Awards | Highlighting Resources in Digital Humanities. Retrieved from http://dhawards.org/

Dixon, J., & Barata, A. C. (1999). American Art Museums on the Web. *INSPEL*, *33*, 209–214.

Donovan, K. (1997). The Best of Intentions: Public Access, the Web & the Evolution of Museum Automation. In D. Bearman & J. Trant (Eds.), *Selected Papers from Museums and the Web 97*. Pittsburgh: Archives & Museum Informatics. Retrieved from http://www.museumsandtheweb.com/mw97/speak/donovan.htm

Dörk, M. (2012). *Visualization for Search: Exploring Complex and Dynamic Information Spaces*. University of Calgary, Calgary. Retrieved from http://mariandoerk.de/phd/

Dörk, M., Carpendale, S., & Williamson, C. (2011). The Information Flaneur: A Fresh Look at Information Seeking. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems* (pp. 1215–1224). Vancouver, BC, Canada: ACM. Retrieved from http://mariandoerk.de/informationflaneur/chi2011.pdf

Driscoll, D. L., Appiah-Yeboah, A., Salib, P., & Rupert, D. J. (2007). Merging Qualitative and Quantitative Data in Mixed Methods Research: How To and Why Not. *Ecological and Environmental Anthropology (University of Georgia)*. Retrieved from http://digitalcommons.unl.edu/icwdmeea/18

Drucker, J. (2011). Humanities Approaches to Graphical Display, 5(1). Retrieved from http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html

Drucker, J. (2013). Performative Materiality and Theoretical Approaches to Interface, 7(1). Retrieved from http://www.digitalhumanities.org/dhq/vol/7/1/000143/000143.html

Dumas, B., Moerman, B., Trullemans, S., & Signer, B. (2014). ArtVis: combining advanced visualisation and tangible interaction for the exploration, analysis and browsing of digital artwork collections (pp. 65–72). ACM Press. http://doi.org/10.1145/2598153.2598159

Fallman, D. (2008). The interaction design research triangle of design practice, design studies, and design exploration. *Design Issues*, *24*(3), 4–18.

Few, S. (2012, May 18). Data Art vs. Data Visualization: Why Does a Distinction Matter? Retrieved from http://www.perceptualedge.com/blog/?p=1245

Finley, K. (2016). The Days of Microsoft Internet Explorer Are Numbered—But Its Sorry Legacy Will Live On. Retrieved 26 June 2016, from http://www.wired.com/2016/01/the-sorry-legacy-of-microsoft-internet-explorer/

Finnis, J., Chan, S., & Clements, R. (2011). *Let's Get Real. How to evaluate success online?* Brighton, UK: WeAreCulture24.

Frayling, C. (1994). Research in Art and Design (Royal College of Art Research Papers, Vol 1, No 1, 1993/4) [Monograph]. Retrieved 2 March 2016, from http://researchonline.rca.ac.uk/384/

Fry, B. (2009). the preservation of favoured traces. Retrieved 3 February 2010, from http://benfry.com/traces/

Friendly, M., & Denis, D. (2001). Milestones in the history of thematic cartography, statistical graphics, and data visualization. Retrieved 18 May 2010, from http://www.datavis.ca/milestones/

Google. (2011). Google Art Project. Retrieved from https://www.google.com/culturalinstitute/

Gordon, S. (1995). Is Anybody Out There? (museums, audiences and the World Wide Web). Retrieved from http://www.archimuse.com/publishing/ichim95_vol2/gordon.pdf

Gorgels, P. (2013). Rijksstudio: Make Your Own Masterpiece! Presented at the Museums and the Web, Portland, Oregon. Retrieved from http://mw2013.museumsandtheweb.com/paper/rijksstudio-make-your-own-masterpiece/

Gorz, G., & Holst, N. (1997). The Digital Behaim Globe (1492). Retrieved from http://www.archimuse.com/publishing/ichim97/gorz.pdf

Grady, J. (2006). Edward Tufte and the Primise of a Visual Social Science. In L. Pauwels (Ed.), *Visual Cultures of Science: Rethinking Representational Practices in Knowledge Building and Science Communication*. UPNE.

Gray, M. (1997). Web Growth Summary. Retrieved from http://www.mit.edu/people/mkgray/net/web-growth-summary.html

Greene, S., Marchionini, G., Plaisant, C., & Shneiderman, B. (2000). Previews and overviews in digital libraries: Designing surrogates to support visual information seeking. *Journal of the American Society for Information Science*, *51*(4), 380–393. http://doi.org/10.1002/(SICI)1097-4571(2000)51:4<380::AID-ASI7>3.0.CO;2-5

Griffin, M. (2012). Responsive Comping: Obtaining Signoff without Mockups. Retrieved 19 May 2016, from http://alistapart.com/article/responsive-comping-obtaining-signoff-with-mockups

Guralnick, R. (1997). Bridging the Gap between Research and Education: The Future of Network Technology in the Sciences and Science Museums. Retrieved from http://www.museumsandtheweb.com/mw97/speak/guralnic.html

Hailpern, J., & Hagedorn, J. (2008). VCode & VData (Version 1.2.1). Retrieved from http://social.cs.uiuc.edu/projects/vcode.html

Hall, E. (2008). *Copy As Interface*. Presented at the Web2.0 Expo, San Francisco, Calif. Retrieved from http://www.slideshare.net/mulegirl/copy-as-interface

Hangal, S., Nagpal, A., & Lam, M. (2012). Effective Browsing and Serendipitous Discovery with an Experience-infused Browser. In Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces (pp. 149–158). New York, NY, USA: ACM. https://doi.org/10.1145/2166966.2166994

Hassenzahl, M., & Tractinsky, N. (2006). User experience – a research agenda. *Behaviour & Information Technology*, *25*(2), 91–97. http://doi.org/10.1080/01449290500330331

Helmes, J., O'Hara, K., Vilar, N., & Taylor, A. (2011). Meerkat and Tuba: Design Alternatives for Randomness, Surprise and Serendipity in Reminiscing. In Human-Computer Interaction – INTERACT 2011 (pp. 376–391). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-23771-3_28

Hempel, S. (2013, May 6). John Snow and cholera: how myth helped secure his place in history. Retrieved 10 February 2016, from http://blog.oup.com/2013/05/john-snow-bicentenary-cholera/

Hertenstein, J. H., Platt, M. B., & Veryzer, R. W. (2013). What Is 'Good Design'?: An Investigation of the Complexity and Structure of Design. *Design Management Journal*, *8*(1), 8–21. http://doi.org/10.1111/dmj.12000

Hewson, C. (2008). Internet-mediated research as an emergent method and its potential role in facilitating mixed methods research. In S. Hesse-Biber & P. Leavy (Eds.), *Handbook of Emergent Methods* (pp. 542–570). Guildford Press.

Hibino, S., & Rundensteiner, E. A. (1996). MMVIS: a multimedia visual information seeking environment for video analysis (pp. 15–16). ACM Press. http://doi.org/10.1145/257089.257099

Hinton, S., & Whitelaw, M. (2010). commonsExplorer. Retrieved 13 October 2015, from http://mtchl.net/commons-explorer/

Hong, J.-K., Takahashi, J., Kusaba, M., & Sugita, S. (1995). An Approach to the Digital Museum – Multimedia Systems for an Ethnology Museum. In *ICHIM, Multimedia Computing and Museums* (pp. 87–95).

Hornbæk, K., & Hertzum, M. (2011). The notion of overview in information visualization. *International Journal of Human-Computer Studies*, *69*(7-8), 509–525. http://doi.org/10.1016/j.ijhcs.2011.02.007

Isenberg, T., Isenberg, P., Jian Chen, Sedlmair, M., & Moller, T. (2013). A Systematic Review on the Practice of Evaluating Visualization. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2818–2827. http://doi.org/10.1109/TVCG.2013.126

Jick, T. D. (1979). Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*, *24*(4), 602–611.

Johnson, E. (2011). The Google Art Project: a new generation of museums on the web? [Blog comment]. Retrieved from http://www.curatorjournal.org/digital/the-google-art-project-a-new-generation-of-museums-on-the-web/#comment-68

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research*, *1*(2), 112–133. http://doi.org/10.1177/1558689806298224

Jones, K. S., & Willett, P. (1997). *Readings in Information Retrieval*. Morgan Kaufmann.

Jones, M. (2009). Data as Seductive Material. Retrieved 13 November 2010, from http://www.slideshare.net/blackbeltjones/data-as-seductive-material-spring-summit-ume-march09

Kahn, P., Lenk, K., & Kasman, M. (1997). Real space and cyberspace: a comparison of museum maps and electronic publication maps. In *ICHIM 97: international conference on hypermedia and interactivity in museums* (pp. 99–113). Retrieved from http://cat.inist.fr/?aModele=afficheN&cpsidt=2278214

Keller, T., & Tergan, S.-O. (2005). Visualizing Knowledge and Information: An Introduction. In S.-O. Tergan & T. Keller (Eds.), *Knowledge and Information Visualization* (Vol. 3426, pp. 1–23). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/index/10.1007/11510154_1

Kenderdine, S. (1997). Diving into shipwrecks: Aquanauts in cyberspace. *Bulletin of the Australian Institute for Maritime Archaeology, The*, *21*, 85.

Kramer-Smyth, J. (2007). ArchivesZ: Visualizing Archival Collections. Retrieved 8 November 2010, from http://archivesz.com/

Kim, J. (2007). Modeling task-based information seeking on the web: Application of information seeking strategy schema. *Proceedings of the American Society for Information Science and Technology*, *44*(1), 1–13. http://doi.org/10.1002/meet.1450440239

Kincheloe, J. L. (2001). Describing the Bricolage: Conceptualizing a New Rigor in Qualitative Research. *Qualitative Inquiry*, *7*(6), 679–692. http://doi.org/10.1177/107780040100700601

Koski, B., & Hinton, E. (2014). The PourOver Book. Retrieved from http://nytimes.github.io/pourover/

Koskinen, I., Mikkonen, J., Eckoldt, K., Hänninen, R., Jiang, J., Schultz, B., … Suri, J. F. (2009). Protosketching: Sketching in Experience Prototyping. Presented at the CHI 2009, Boston MA, USA: ACM. Retrieved from http://www2.uiah.fi/~ikoskine/recentpapers/submissions08/protosketching-submission-not-anon.pdf

Kowalski, G. (2010). *Information Retrieval Architecture and Algorithms*. Springer Science & Business Media.

Kräutli, F. (2016). *Visualising Cultural Data. Exploring and Understanding Cultural Data Through Interactive Timelines*. Royal College of Art, London.

Krikelas, J. (1983). Information-Seeking Behavior: Patterns and Concepts. *Drexel Library Quarterly*, *19*(2), 5–20.

Kuhlthau, C. (1991). Inside the Search Process: Information Seeking from the User's Perspective – ProQuest. Retrieved 20 June 2016, from http://search.proquest.com/openview/d6f2fedf267c5f277d70414812bc4d34/1?pq-origsite=gscholar&cbl=1818555

Lamping, J., Rao, R., & Pirolli, P. (1995). A focus+ context technique based on hyperbolic geometry for visualizing large hierarchies. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 401–408). ACM Press/Addison-Wesley Publishing Co. Retrieved from http://dl.acm.org/citation.cfm?id=223956

Large, A., Tedd, L. A., & Hartley, R. J. (1999). *Information Seeking in the Online Age: Principles and Practice*. Bowker-Saur.

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P. O. S., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach (p. 719). ACM Press. http://doi.org/10.1145/1518701.1518813

Leckie, G. J., Pettigrew, K. E., & Sylvain, C. (1996). Modeling the Information Seeking of Professionals: A General Model Derived from Research on Engineers, Health Care Professionals, and Lawyers. *The Library Quarterly: Information, Community, Policy*, *66*(2), 161–193.

LeGrady, G. (2005). Making Visible the Invisible. Retrieved 8 November 2010, from http://www.mat.ucsb.edu/~g.legrady/glWeb/Projects/spl/spl.html

MacDonald, C. (2015). Assessing the user experience (UX) of online museum collections: Perspectives from design and museum professionals. Presented at the MW2015: Museums and the Web 2015, Chicago. Retrieved from http://mw2015.museumsandtheweb.com/paper/assessing-the-user-experience-ux-of-online-museum-collections-perspectives-from-design-and-museum-professionals/

Mäkelä, A. M., & Nimkulrat, N. (2011). Reflection and Documentation in Practice-led Design Research. *Nordes*, *0*(4). Retrieved from http://www.nordes.org/opj/index.php/n13/article/view/98

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. New York: Cambridge University Press. Retrieved from http://nlp.stanford.edu/IR-book/

Mannoni, B. (1997). A virtual museum. *Communications of the ACM*, *40*(9), 61–62. http://doi.org/10.1145/260750.260772

Manovich, L. (2007). Cultural Analytics: Analysis and Visualization of Large Cultural Datasets. Retrieved from http://www.manovich.net/cultural_analytics.pdf

Mansfield, T., Winter, C., Griffith, C., Dockerty, A., & Brown T. (2014). *Innovation Study: Challenges and Opportunities for Australia's Galleries, Libraries, Archives and Museums*. Australian Centre for Broadband Innovation, CSIRO and Smart Services Co-operative Research Centre.

Marchionini, G. (1997). *Information Seeking in Electronic Environments*. Cambridge University Press.

Marcotte, E. (2010). Responsive Web Design. Retrieved 23 May 2016, from http://alistapart.com/article/responsive-web-design

Marty, P. F. (2007). Museum websites and museum visitors: Before and after the museum visit. *Museum Management and Curatorship*, *22*(4), 337–360.

Marty, P. F. (2011). My lost museum: User expectations and motivations for creating personal digital collections on museum websites. *Library & Information Science Research*, *33*(3), 211–219. http://doi.org/10.1016/j.lisr.2010.11.003

Marty, P. F., & Burton Jones, K. (2012). *Museum Informatics: People, Information, and Technology in Museums*. Routledge.

Matthews, R. (1999). The Evolution Of A Multimedia Product Over A Decade: the 'Micro Gallery' family tree – past, present and futures. Retrieved from http://www.archimuse.com/publishing/ichim99/matthews.pdf

Mayo, K., & Newcomb, P. (2009). How the Web was won. An oral history of the internet. Retrieved 29 June 2016, from http://www.vanityfair.com/news/2008/07/internet200807?currentPage=8

McRaney, D. (2011). The Sunk Cost Fallacy. Retrieved from https://youarenotsosmart.com/2011/03/25/the-sunk-cost-fallacy/

MDN. (2016). CSS3. Retrieved 26 June 2016, from https://developer.mozilla.org/en-US/docs/Web/CSS/CSS3

M.L.I.S, S. J. T. (2010). Website Statistics 2.0: Using Google Analytics to Measure Library Website Effectiveness. *Technical Services Quarterly*, *27*(3), 261–278. http://doi.org/10.1080/07317131003765910

Moere, A. V., & Purchase, H. (2011). On the role of design in information visualization. *Information Visualization*, *10*(4), 356–371. http://doi.org/10.1177/1473871611415996

Molineux, K. (2014). Permanent Collection Displays. In B. Lord & M. Piacente (Eds.), *Manual of Museum Exhibitions*. Rowman & Littlefield.

Mooers, C. N. (1951). Zatocoding applied to mechanical organization of knowledge. *American Documentation*, *2*(1), 20–32. http://doi.org/10.1002/asi.5090020107

Morrison, A. (1995). The Micro Gallery – Observations From Three Projects: London; San Diego; Washington D.C. In *ICHIM, Hands on: Hypermedia and Interactivity in Museums* (pp. 13–20). Retrieved from http://www.archimuse.com/publishing/ichim95_vol2/morrison.pdf

Musalem, A. M. (1995). A Multimedial Database System – Managing a Virtual Collection of Art and Architectural Works. In *ICHIM, Multimedia Computing and Museums* (pp. 39–56).

MW. (1997). Best Research Site. Retrieved from http://www.museumsandtheweb.com/mw97/mw97brs.html

MW. (1998). Best Research Site. Retrieved from http://www.museumsandtheweb.com/mw98/best/win_research.html

Naisbitt, J. (1984). *Megatrends: Ten New Directions Transforming Our Lives*. Warner Books.

Naumann, P. (2012). Art in detail: the Google Art Project. *Artonview*, (70). Retrieved from https://issuu.com/nationalgalleryofaustralia/docs/artonview_70

Neish, P. (2014). *Linked Data: thinking big, starting small*. Retrieved from http://peter.neish.net/wp-content/uploads/sites/6/2014/02/VALA2014-Session-10-Neish-Paper.pdf

NGA. (2015). *2015 Corporate Plan*. Canberra: National Gallery of Australia.

Nielsen, J. (1993). Iterative Design of User Interfaces. Retrieved 10 May 2016, from https://www.nngroup.com/articles/iterative-design/

Nilsson, T. (1997). The Interface of a Museum: Text, Context and Hypertext in a Performance Setting. Retrieved from http://www.archimuse.com/publishing/ichim97/nilsson.pdf

Norman, D., & Nielsen, J. (n.d). The Definition of User Experience (UX). Retrieved 24 May 2016, from https://www.nngroup.com/articles/definition-user-experience/

Oates, G. (2015). The V&A Spelunker. Retrieved from http://www.vam.ac.uk/blog/digital-media/the-va-spelunker

Oberoi, S., Stein, R., & Arnold, K. (2015). New architectures for online collections and digitization | MW2015: Museums and the Web 2015. Presented at the MW2015: Museums and the Web 2015, Chicago. Retrieved from http://mw2015.museumsandtheweb.com/paper/new-architectures-for-online-collections-and-digitization/

Oddy, R. N., Belkin, N. J., & Brooks, H. M. (1982). ASK for information retrieval: Part I. Background and theory. *Journal of Documentation*, *38*(2), 61–71.

OUP. (2013). *Oxford Dictionary of English*. Oxford: Oxford University Press.

Peacock, D., Ellis, D., & Doolan, J. (2004). Searching For Meaning: Not Just Records. In D. Bearman & J. Trant (Eds.), *Museums and the Web 2004: Proceedings*. San Francisco, Calif.: Archives & Museum Informatics. Retrieved from http://www.museumsandtheweb.com/mw2004/papers/peacock/peacock.html

Penman, R. (2015). Let Links Be Links. Retrieved 4 June 2016, from http://alistapart.com/article/let-links-be-links

Piacente, M. (1996). Surf's up : museums and the World Wide Web.

Pickard, J. (2006). A Standardista's Alphabet. Retrieved 26 June 2016, from http://alistapart.com/article/alphabet

Pierroux, P. (1998). *Art in Networks: Information and Communciation Technology in Art Museums*. University of Oslo, Oslo, Sweden. Retrieved from http://folk.uio.no/palmyre/Hovedoppgave/index.html

Pohl, M., Wiltner, S., & Miksch, S. (2010). Exploring information visualization: describing different interaction patterns. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond time and errors: novel evaLuation methods for Information Visualization* (pp. 16–23). Retrieved from http://dl.acm.org/citation.cfm?id=2110195

Proctor, N. (2011). The Google Art Project: A New Generation of Museums on the Web? *Curator: The Museum Journal*, *54*(2), 215–221. http://doi.org/10.1111/j.2151-6952.2011.00083.x

Proença, A., Brito, M., Ramalho, T., & Regalo, H. (1998). Using the Web to give life to Museums. Presented at the Museums and the Web. Retrieved from http://www.museumsandtheweb.com/mw98/papers/proenca/proenca_paper.html

Rao, R., & Card, S. K. (1994). The table lens: merging graphical and symbolic representations in an interactive focus+ context visualization for tabular information. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 318–322). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=191776

Rayward, W. B., & Twidale, M. B. (1999). From Docent to Cyberdocent: Education and Guidance in the Virtual Museum. *Archives and Museums Informatics*, *13*, 23–53.

Reimer, L. (2011, June 22). Following A Web Design Process. Retrieved 19 May 2016, from https://www.smashingmagazine.com/2011/06/following-a-web-design-process/

Rice, R. E., McCreadie, M., & Chang, S.-J. L. (2001). *Accessing and Browsing Information and Communication*. MIT Press.

Rijksmuseum. (2012). Rijks Studio. Retrieved 1 July 2016, from https://www.rijksmuseum.nl/en

RijksMuseum. (2016). Redesigned Rijksstudio: ready for the mobile era – News – What's on. Retrieved 29 June 2016, from https://www.rijksmuseum.nl/en/whats-on/news/redesigned-rijksstudio-ready-for-the-mobile-era

Robbins, J. N. (2012). *Learning Web Design: A Beginner's Guide to HTML, CSS, JavaScript, and Web Graphics*. O'Reilly Media, Inc.

Rodenbeck, E. (2008). Information Visualization is a Medium. Presented at the Emerging Technology, San Diego, California. Retrieved from http://en.oreilly.com/et2008/public/schedule/detail/1588

Rogers, Y. (2004). New theoretical approaches for human-computer interaction. *Annual Review of Information Science and Technology*, *38*(1), 87–143. http://doi.org/10.1002/aris.1440380103

Roser, M. (2014). Number of internet users by world region, 1990-2014. Retrieved 28 June 2016, from https://ourworldindata.org/grapher/number-of-internet-users-by-world-region-1990-2014

Rust, C., Mottram, J., & Till, J. (2007). *AHRC Review of Practice-Led Research in Art Design & Architecture*. Arts and Humanities Research Council, Sheffield Hallam University, Nottingham Trent University & University of Sheffield. Retrieved from http://archive.org/details/ReviewOfPractice-ledResearchInArtDesignArchitecture

Sanderson, M., & Croft, W. B. (2012). The History of Information Retrieval Research. *Proceedings of the IEEE*, *100*(Special Centennial Issue), 1444–1451. http://doi.org/10.1109/JPROC.2012.2189916

Saracevic, T., Kantor, P., Chamis, A. Y., & Trivison, D. (1988). A Study of Information Seeking and Retrieving. I. Background and Methodology. *Journal of the American Society for Information Science*, *39*(3), 161–176.

Schön, D. A. (2008). *The Reflective Practitioner: How Professionals Think In Action*. Basic Books.

Seligman, M. E. P., & Csikszentmihalyi, M. (2014). Positive Psychology: An Introduction. In *Flow and the Foundations of Positive Psychology* (pp. 279–298). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-94-017-9088-8_18

SFMOMA. (2009). ArtScope. Retrieved 8 April 2010, from http://www.sfmoma.org/projects/artscope/

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*. Retrieved from http://portal.acm.org/citation.cfm?id=834354

Shneiderman, B., & Plaisant, C. (2006). Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies (p. 1). ACM Press. http://doi.org/10.1145/1168149.1168158

Skov, M., & Ingwersen, P. (2014). Museum Web search behavior of special interest visitors. *Library & Information Science Research*, *36*(2), 91–98. http://doi.org/10.1016/j.lisr.2013.11.004

Smith, P. F. (2003). *The dynamics of delight: architecture and aesthetics*. Routledge.

Spool, J. M. (2014). Code Sketching – A Stretch Goal for Your Design Superpower. Retrieved from https://articles.uie.com/sketching_in_code/

Stamen Design. (n.d). SFMoMA. Retrieved 28 June 2016, from http://stamen.com/work/sfmoma/

Stasko, J., & Zhang, E. (2000). Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *IEEE Symposium on Information Visualization, 2000. InfoVis 2000* (pp. 57–65). http://doi.org/10.1109/INFVIS.2000.885091

StatCounter. (2016). Global Stats: Screen Resolution – August 2011 to June 2016. Retrieved from http://gs.statcounter.com/#resolution-ww-monthly-201108-201606

Steele, J., & Iliinsky, N. (2010). *Beautiful Visualization: Looking at Data through the Eyes of Experts*. O'Reilly Media, Inc.

Sula, C. A. (2013). Quantifying Culture: Four Types of Value in Visualization. In *Electronic Visualisation in Arts and Culture*. Retrieved from http://chrisalensula.org/quantifying-culture-four-types-of-value-in-visualization/

Sullivan, G. (2006). Research Acts in Art Practice. *Studies in Art Education*, *48*(1), 19–35.

Talagala, N., Asami, S., Patterson, D., Futernick, B., & Hart, D. (1998). The Berkeley-San Francisco Fine Arts Image Database. Retrieved from https://www.researchgate.net/publication/2876510_The_Berkeley-San_Francisco_Fine_Arts_Image_Database

Taramigkou, M., Bothos, E., Apostolou, D., & Mentzas, G. (2013). Fostering serendipity in online information systems. In 2013 International Conference on Engineering, Technology and Innovation (ICE) IEEE International Technology Management Conference (pp. 1–10). https://doi.org/10.1109/ITMC.2013.7352707

Tasich, T., & Villaespesa, E. (2013). Meeting the Real User: Evaluating the Usability of Tate's Website | MW2013: Museums and the Web 2013. Presented at the Museums and the Web, Portland, Oregon. Retrieved from http://mw2013.museumsandtheweb.com/paper/meeting-the-real-user-evaluating-the-usability-of-tates-website/

Taylor, R. S. (1962). The process of asking questions. *American Documentation*, *13*(4), 391–396. http://doi.org/10.1002/asi.5090130405

Teather, L., & Wilhelm, K. (1999). 'Web Musing': Evaluating Museums on the Web from Learning Theory to Methodology. Presented at the Museums and the Web, Boston MA, USA: Archives & Museum Informatics. Retrieved from http://www.archimuse.com/mw99/papers/teather/teather.html

TOMS, E. G. (2000). Understanding and facilitating the browsing of electronic text. *International Journal of Human-Computer Studies*, *52*(3), 423–452. http://doi.org/10.1006/ijhc.1999.0345

Toney, S. R., & Donoghue, K. (1998). New Web-Based Interfaces to Old Databases. Presented at the Musuems and the Web, Toronto. Retrieved from http://www.museumsandtheweb.com/mw98/papers/toney/toney_paper.html

Tufte, E. R. (2006). *Beautiful evidence*. Cheshire, Conn: Graphics Press.

Turnbull, D. (2000). Augmenting Information Seeking on the World Wide Web Using Collaborative Filtering Techniques. Canada: University of Toronto. Retrieved from https://www.ischool.utexas.edu/~donturn/research/augmentis.html

Unger, R., & Chandler, C. (2012). *A Project Guide to UX Design: For user experience designers in the field or in the making*. New Riders.

Vande Moere, A. (2007). Aesthetic Data Visualization as a Resource for Educating Creative Design. In *Proceedings of the 12th International CAAD Futures Conference*. Sydney, Australia: Springer.

Verba, D. (2008). Sketching in Code: the Magic of Prototyping. Retrieved 10 May 2016, from http://alistapart.com/article/sketchingincode

Victoria & Albert Museum. (2011, June 28). Ceramics Phase 2, The Ceramics Study Galleries. Retrieved 15 March 2016, from http://www.vam.ac.uk/content/articles/d/ceramics-study-galleries/

Villaespesa, E., & Tasich, T. (2012). Making Sense of Numbers: A Journey of Spreading the Analytics Culture at Tate. Presented at the Museums and the Web, San Diego, California. Retrieved from http://www.museumsandtheweb.com/mw2012/programs/making_sense_of_numbers_a_journey_of_spreadi.html

Walker Art Center. (2011). Walker Art Center Collections. Retrieved 6 June 2015, from http://www.walkerart.org/collections/browse

W3C. (2015a). About W3C. Retrieved from http://www.w3.org/Consortium/

W3C. (2015b). History of the Web. Retrieved 19 November 2015, from http://webfoundation.org/about/vision/history-of-the-web/

Watz, M. (2012). Screen Saver Culture. Retrieved 4 May 2016, from https://www.scribd.com/doc/78375321/20120116-screen-saver-culture

Webster, F. (2014). *Theories of the Information Society*. Routledge.

Weiler, A. (2005). Information-Seeking Behavior in Generation Y Students: Motivation, Critical Thinking, and Learning Theory. *The Journal of Academic Librarianship*, *31*(1), 46–53. http://doi.org/10.1016/j.acalib.2004.09.009

Whitelaw, M. (2009). The Visible Archive. Retrieved from http://visiblearchive.blogspot.com/

Whitelaw, M. (2012a). Towards Generous Interfaces for Archival Collections. Presented at the International Council on Archives, Brisbane. Retrieved from http://www.ica2012.com/files/data/Full%20papers%20upload/ica12Final00423.pdf

Whitelaw, M. (2012b). Manly Images. Retrieved from http://mtchl.net/manly-images/

Whitelaw, M. (2015). Generous Interfaces for Digital Cultural Collections, *9*(1). Retrieved from http://www.digitalhumanities.org/dhq/vol/9/1/000205/000205.html

Whitelaw, M., & Hinton, S. (2010). commonsExplorer. Retrieved 8 November 2010, from http://creative.canberra.edu.au/cex/

White, R. W., Kules, B., Drucker, S. M., & others. (2006). Supporting exploratory search, introduction, special issue, communications of the ACM. *Communications of the ACM*, *49*(4), 36–39.

Wolfe, G. (1994). The (Second Phase of the) Revolution Has Begun. Retrieved 9 March 2016, from http://www.wired.com/1994/10/mosaic/

Wroblewski, L. (2009, March 29). Defining User Experience. Retrieved 24 May 2016, from http://www.lukew.com/ff/entry.asp?797

Yee, J. S. R. (2010). Methodological Innovation in Practice-Based Design Doctorates. *Journal of Research Practice*, *6*(2), 15.

Zhang, J., & Marchionini, G. (2005). Evaluation and evolution of a browse and search interface: Relation Browser. Presented at the In: Proceedings of the 2005 national conference on Digital government research. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/ions;jsessionid=BF09AC25F80B0111946150E982C236F0?doi=10.1.1.116.5149

Zhang, Y. (2000). Using the Internet for survey research: A case study. *Journal of the American Society for Information Science*, *51*(1), 57–68. http://doi.org/10.1002/(SICI)1097-4571(2000)51:1<57::AID-ASI9>3.0.CO;2-W